Supported Employment Demonstration Final Impact and Cost-Benefit Analysis Report

Contract #SS00-16-60014
Contracting Officer's Representative (COR):
Marion McCoy, PhD
Alternate COR: Ty Turner, MA

April 2023

Authors
Jeffrey A. Taylor, PhD
Mustafa Karakus, PhD
Jarnee Riley, MS
David Salkever, PhD
William Frey, PhD
Howard Goldman MD, PhD
Jocelyn Marrow, PhD

Submitted to:
Social Security Administration
Office of Acquisition and Grants
Division of Programs Contracts
1540 Robert M. Ball Building
6401 Security Blvd
Baltimore, MD 21235

Submitted by:
Westat
An Employee-Owned Research Corporation®
1600 Research Boulevard
Rockville, MD 20850-3129
(301) 251-1500

Table of Contents

Table of Contents (continued)

Table of Contents (continued)

Table of Contents (continued)

Tables

Table of Contents (continued)

Tables

Exhibits

List of Acronyms

ACA	Affordable Care Act
AHRQ	Agency for Healthcare Research and Quality
BLS	Bureau of Labor Statistics
CARES Act	Coronavirus Aid, Relief, and Economic Security Act
CHAID	Chi-squared Automatic Interaction Detection
CDC	Centers for Disease Control and Prevention
COR	Contracting Officer's Representative
CSI	Colorado Symptom Index
GED	General Education Diploma
HHS	Health and Human Services
FTE	Full-Time Equivalent
IPS	Individual Placement and Support
MCF	Mental Computation Fluency
MCS	Mental Component Score
MHTS	Mental Health Treatment Study
MIS	Management Information System
MMS	Medication Management Support
NCC	Nurse Care Coordinator
QCEW	Quarterly Census of Employment and Wages
PCF	Primary Care First
PCS	Physical Component Score
RA	Research Assistant
RCT	Randomized Controlled Trial
RFP	Request for Proposal
SE	Supported Employment
SED	Supported Employment Demonstration
SGA	Substantial Gainful Activity
SNAP	Supplemental Nutrition Assistance Program
SSA	Social Security Administration
SSDI	Social Security Disability Insurance
SSI	Supplemental Security Income
TA/QA	Technical Assistance/Quality Assurance Specialist
TANF	Temporary Assistance for Needy Families

Executive Summary

Overview

The Supported Employment Demonstration (SED) aims to improve employment outcomes for Social Security disability applicants alleging a mental impairment who were recently denied benefits. The Social Security Administration (SSA) seeks to answer whether offering the Individual Placement and Support (IPS) model of employment services and behavioral health and other services fosters employment and clinical recovery that leads to self-sufficiency, improved quality of life, and less demand for disability benefits.

The SED used a randomized controlled trial (RCT) experimental design to test two treatment alternatives, Full-Service and Basic-Service, against a Usual Services (control) group. The SED enrolled and randomized 3,000 participants, aged 18-49 years, to one of the three study arms across 30 demonstration sites in 20 states for 3 years of study participation. The treatment interventions integrated supported employment (SE) with behavioral health treatment following the evidence-based IPS model of employment services. The interventions provided to both treatment groups also included the provision of care management services to address barriers to employment and modest financial support for individual work-related expenses and out-of-pocket expenses associated with behavioral health and other care management services not covered by health insurance. The Full-Service intervention also included Medication Management Support (MMS) delivered by a Nurse Care Coordinator (NCC). Usual Services participants received a comprehensive resource manual and sought out services independently as they normally would.

Implementation

Westat selected 20 “large” sites for recruiting 120 participants each and 10 “small” sites for recruiting 60 participants each to yield the target sample size of 3,000 study participants. Sites represented typical community agencies that serve people with mental illness and other low-income populations. Twenty-five of the 30 agencies belonged to the IPS Learning Community, a national network offering the IPS model of employment services. All agencies offered care manager and benefits counseling services, and most offered nursing services for medication delivery or management.

Staffing the SED involved establishing 60 treatment teams—two per agency—including a team lead, IPS specialist, care manager, and NCC (for Full-Service teams only). Westat also assigned a technical assistance/quality assurance (TA/QA) specialist to each agency. SED staff received training on the study protocol and position-relevant aspects via in-person trainings, taped webinars, conference calls, and one-on-one sessions.

The SED recruitment process included several steps: initial contact and eligibility screening; recruitment information meetings; competency screening and informed consent; baseline interview; randomization; and wrap-up activities. Two Westat field directors oversaw and managed the study’s recruitment team, which included three regional field supervisors; 30 local recruiters or research assistants (RAs); and 13 traveling recruiters. Recruitment staff completed pre- and post-classroom home study and classroom training. The study also established procedures to pay for uninsured participants' healthcare until they could apply for state or Affordable Care Act (ACA) coverage.

Enrollment rates for the SED ranged from 23.0 percent to 26.2 percent. The lower end of 23 percent included the 3,000 denied applicants who enrolled in the study within the context of the larger population of denied applicants who were eligible for the study and contacted for recruitment. However, although we sent letters and attempted to make personal contact with all those who were eligible, we were not able to reach some and have no way of knowing that they ever personally became aware of the study and their eligibility. Hence, the higher end of 26.2 percent enrollment rate includes the 3,000 enrollments within the context of only those eligible denied applicants for whom we have confirmation that they at least heard about the study and knew they were eligible.

After enrollment, we learned that 56 study participants were receiving benefits from another previous application. These applicants denied disability on their "current" application were deemed ineligible for the study given they were receiving benefits from a "previous" application. While the study continued to serve these enrollees, we removed them from all subsequent analyses since they were technically ineligible, thereby reducing the effective sample size from 3,000 to 2,944. Random assignment of the 2,944 eligible study participants resulted in 976 (33.2%) study participants assigned to the Full-Service treatment group; 987 (33.5%) assigned to the Basic-Service treatment group; and 981 (33.3%) assigned to the Usual Services (control) group. Among the eligible study participants, 83 died during the 3-year study enrollment period (20 Full-Service, 30 Basic-Service, and 33 Usual Services). The death rate was not significantly different across the study arms.

Participation in Research

In addition to completing a baseline interview at enrollment, we asked SED study participants to complete 12 quarterly interviews throughout their 3-year enrollment period to measure key outcomes and assess impacts of the interventions. We measured the impact of the Full-Service and Basic-Service interventions on outcomes in five domains: employment and earnings, SSA benefit receipt, health status, quality of life, and healthcare utilization. We used data from the quarterly interviews that rely on participant response and recall, measuring each of these outcomes in the domains of employment and earnings, health status, quality of life, and healthcare utilization. We used data outside of the participant surveys to assess outcomes related to SSA benefit receipt.

The study team encountered challenges in locating and contacting study participants to collect the quarterly interview data. Response rates among eligible participants held above 70 percent for the first 2 years of study enrollment (quarters 1-8). The third year of the study saw a drop-off in completion rates; by quarter 12, roughly two-thirds (65.3%) of eligible enrollees completed the survey.

The study team limited the analysis of outcomes related to employment and healthcare utilization to those participants who completed enough quarterly surveys to provide an accurate accounting of their work and health over the course of the study. Similarly, we limited analysis of health status and quality of life outcomes reported in the annual interviews (quarters 4, 8, and 12) to participants who completed those specific quarterly interviews. When estimating impacts for those outcome measures that rely on participant response and recall from the follow-up interviews, the study team followed three approaches to test and adjust for attrition patterns: nonresponse weighting, regression adjustment, and bounding. SSA administrative records provide complete data for all SED participants for outcomes related to receipt of SSA benefits.

Impact Evaluation

Employment and Earnings

Employment impacts of the program are significant and increasing over time.

Selected Intervention Outcomes

Significantly better than Usual Services

Better but not significantly different than Usual Services

Outcome	Full-Service	Basic-Service	Usual Services
Employment rate (employed any time in 3-year study period)	74%	74%	64%
Employment rate in Year 1	54%	58%	50%
Employment rate in Year 2	60%	58%	49%
Employment rate in Year 3	54%	53%	43%
Total earnings over 3-year study period (all participants)	$17,925	$17,556	$13,547
Total earnings over 3-year study period (had at least one job)	$24,216	$23,588	$21,123

- Full-Service and Basic-Service participants had significantly higher rates of employment than the Usual Services participants during each year of study enrollment. Compared to those in the Usual Services group, those with Basic-Service had annual employment rates that were respectively 6 percent, 7 percent and 9 percent higher over Years 1, 2 and 3 of the program.[1] Those with Full-Service had rates that were 4percent, 11 percent, and 11 percent higher over Years 1, 2, and 3, respectively. There were no significant differences in employment rates between the Full-Service and Basic-Service participants.

- The Full-Service and Basic-Service interventions positively impacted total earnings during study enrollment. In Year 1, there were no significant differences in total annual income between the three treatment groups; however, in Year 2, participants in the Full-Service group had on average $1,146 more total income than participants in the Usual Services group. In Year 3, participants in the Full-Service group earned on average $2,725 more than the Usual Services group and participants in the Basic-Service group earned $1,978 more total income than those in the Usual Services group. There were no significant differences in earnings between the Full-Service and Basic-Service groups.

- The Full-Service participants were significantly more likely than Usual Services participants to achieve substantial gainful activity (SGA) by the third year of the program. Participants in the Full-Service group who were interviewed in quarter 12 had SGA achievement rates nearly 3 percentage points higher than those in the Usual Services group. Basic-Service alone had an impact that was positive but not statistically significant. There were no significant differences in SGA achievement rates between the Full-Service and Basic-Service groups.

---

[1] All differences are regression-adjusted impacts, which means they are more likely to be true differences attributed to program treatments, since other variables that might impact the outcomes are being held constant.

- The Full-Service and Basic-Service participants worked significantly more hours than the Usual Services participants in the third year of the program. In Years 1 and 2, there were no significant differences between total hours worked between the two treatment groups; however, in Year 3, participants in the Full-Service group worked an average of 170 hours more—and participants in the Basic-Service group worked 130 hours more—than participants in the Usual Services group. The same trend was seen when measuring total annual weeks employed. In Year 3 those in Full-Service worked 5.6 weeks more, and those in Basic-Services worked 4.3 weeks more, compared to those in Usual Services. There were no significant differences in hours worked or weeks worked between the Full-Service and Basic-Service participants.

- Basic-Service and Full-Service participants who worked at least one job earned more annual income than Usual Services participants in Year 3. Earnings for those who worked at least one job were higher among those in both Full-Service and Basic-Service groups compared to the Usual Services group, but the differences were only significant in Year 3 ($2,588 more for Full-Service participants and $1,993 for Basic-Service participants).

- Basic-Service and Full-Service participants who worked at least one job were more likely to work steadily (at least half-time) compared to Usual Services participants in Year 3. In Years 1 and 2, the steady worker percentages were remarkably similar across the three study groups. However, in Year 3, the percent of steady workers was 11 percentage points higher in the Full-Service group and 8.5 points higher in the Basic-Services group compared to the Usual Services group. There were no significant differences between the Full-Service and Basic-Service participants.

Predictors of Employment

Recipients of Full- or Basic-Services are more likely to find employment. Among those recipients, the employment gains for Black, Hispanic, multiracial, less than college-educated and male participants have been greater than for White non-Hispanic, more educated, and female participants, respectively.

- Predictors of employment included: participation in Full- or Basic-Services treatment group, having better mental health, having better physical health, and having fewer functional limitations. After controlling for other factors, the marginal effect of being in the Full-Service group was a 9-point increase in employment during the 36-month study period. Participation in the Basic-Services group resulted in an 8-point marginal benefit to employment during the study period. Other smaller but significant predictors of employment included: having better mental and physical health (+0.3% each, respectively), better upper body function (+0.9%), and better *community mobility* (+3%).

- While people of all races benefited from participation in an intervention group, the gains in employment rates were highest for Black, Hispanic, and multiracial participants. On average over the 3-year study, the employment rate of Black participants was 13.7 percent higher in the Full-Service program and 8.9 percent higher in the Basic-Service program compared to the rate for Black participants in the Usual Services group. White participants in the Full- and Basic-Services groups averaged a 4 percent and 1.7 percent increase, respectively, over White participants in the Usual Services group. Hispanic and multiracial participants benefited the most from SED participation. Hispanics in the Full- or Basic-Service groups had employment rates 17.1 percent and 15.4 percent higher (respectively) than Hispanic in the Usual Services group. Multiracial participants in the Full- and Basic-Service groups had employment rates that were 17.6 percent and 28.8 percent higher (respectively) than multiracial participants in the Usual Services group.

- Participants with 12 years of education or less experienced the largest employment gains as a result of participation in SED. On average, the employment rate of participants with less than a high school education was 11.6 percent higher in the Full-Service group compared to those with less than a high school education in the Usual Services group. Participants with a high school degree also benefited significantly by participating in the Full- or Basic-Service group, with employment rates 11.9 percent and 11.5 percent higher, respectively. Participants with some college or technical college also benefited but to a lesser extent. There were not significant benefits for those with an associate or bachelor's degree or higher in either the Full- or Basic-Service group.

- Men experienced the largest employment gains as a result of SED participation. On average, the employment rate of male participants was 12.6 percent higher in the Full-Service group and 10.9 percent higher in the Basic-Service group compared to males in the Usual Services group. For women, those gains were smaller, and only significant for the Full-Service group (6.4% higher than females in the Usual Services group).

Earnings Impacts

Earnings impacts of the program are significant and increasing over time.

- Participation in the Full- and Basic-Service groups resulted in increased earnings for most demographic groups. However, males, Black and multiracial participants, those with an associate degree, and those working at the time of enrollment experienced the largest increases in earnings over the study period. Males in the Full-Service group on average earned $8,426 more, and males in the Basic-Service group earned $7,102 more, than males in the Usual Services group. The differences between service groups for females were positive but small and not significant. Black participants in the Full-Service group on average earned $6,195 more—and Black participants in the Basic-Service group earned $5,043 more—than Black participants in the Usual Services group. Differences between service groups for White participants were positive but not significant. Multiracial participants in the Basic-Service group averaged $7,240 more in earnings than multiracial participants in the Usual Services group. The differences for multiracial participants in the Full-Service group were large and positive but not statistically significant. Finally, participants with an associate degree in the Full-Service and Basic-Service groups showed significantly more income on average (+$12,927, +$14,186, respectively) compared to those with an associate degree in the Usual Services group.

- Late enrollees experienced larger gains in earnings compared to early enrollees. Participants in the combined treatment group who enrolled in quartiles 3 and 4, on average, saw larger earnings gains compared to those in the combined treatment group who enrolled in quartiles 1 and 2. The first quartile of treatment group enrollees (enrolled from December 5, 2017 through March 5, 2018) did not see significant impacts on earnings relative to the control group in their first, second, or third year of enrollment. A possible explanation for the larger earnings impacts among later enrollees compared to early enrollees is that sites learned over time how to better serve SED participants. Another explanation factor effecting early enrollees was the time between the disability denial decision and enrollment. This delay resulted in a median difference of 116 days between SSA denial and enrollment in the SED for quartile 1, compared to 56, 42, and 37 days for quartiles 2, 3, and 4, respectively. The longer delay may have meant that some participants in quartile 1 were too far removed from the denial decision to benefit from the intervention.

Appeals and Other Measures of SSA Benefits

The SED program had fewer benefits for participants in the areas of disability enrollment, healthcare utilization, mental health, and quality of life.

Selected Intervention Outcomes

Significantly better than Usual Services

Better but not significantly different than Usual Services

Outcome	Full-Service	Basic-Service	Usual Service
Appeal attempts during 3-year study period	44.6%	46.0%	45.6%
Disability payments during 3-year study period	$27,399	$30,749	$30,152
Colorado Symptom Index 3-year difference (lower is better)	-6.6	-5.5	-6.4
Mental Health (SF-12 MCS) 3-year difference (higher is better)	5.6	5.0	5.6
Physical Health (SF 12 PCS) 3-year difference (higher is better)	2.5	1.5	2.1
Average number of routine mental health visits (year 3)	1.3	1.2	1.3
Average number of routine employment support visits (year 3)	.42	.44	.22
Average healthcare utilization costs over 3-year study period	$10,277	$9,378	$10,873
Hospital stays for a mental health problem	.04	.07	.06

- **The Full-Service and Basic-Service treatment did not have a significant impact on appeal attempts or disability payments.** Among those enrolled in the SED, approximately 15 percent (439 participants) received an allowance for disability benefits during the 3 years of enrollment in the study. Almost half (45-46%) of participants filed at least one appeal during the study period. Among those approved for disability, the average total disability payments made during the study was approximately $30,000 per approved participant. The treatment did not have a significant impact on these average amounts.

- **The Full-Service and Basic-Service treatment did not have a significant impact on the mental health, quality of life, or number of arrests of participants.** On average, study participants showed significant improvement in mental health each year of the study. However, when comparing the study arms, the improvements in mental health status were not significantly larger for Full-Service or Basic-Service participants compared to Usual Services. The average quality-of-life increases over the course of the study are generally around a half point on the scale. There were no significant differences in the changes in quality-of-life scores at each annual survey over the course of the study across the study arms. The number of arrests was only significantly lower in year 2 for the Basic-Services group (-0.06), but the effect did not persist over Years 2 and 3.

- **Full-Service participants had a small but significant reduction in hospital stays for mental health problems compared to Usual Services participants.** For Full-Service participants, there were .03 fewer hospital stays for mental health problems (over the entire study period) compared to those who received Usual Services. There were no similar differences for Basic-Services. Nor were there any reductions in emergency room (ER) visits for mental health problems relative to Usual Services. Similarly, there were no significant impacts of the Full-Service or the Basic-Service interventions on the numbers of routine mental health or general health visits each year.

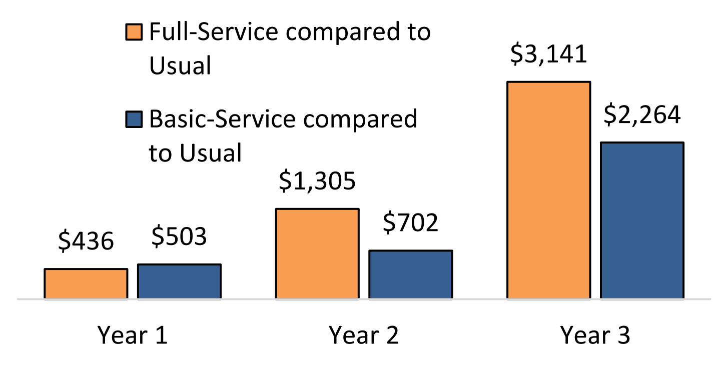
- **Participants in the Basic-Service group were more likely than the Usual Services participants to complete preventive care visits during their enrollment in the study.** On average, Basic-Service participants had 0.07 more preventive care visits (over the entire study period) compared to those who received Usual Services. Participants in the Full-Service group also showed an increased likelihood to complete preventive care, but the difference was not significant.

- **There were no significant impacts of the Full-Service or the Basic-Service interventions on the numbers of routine mental health or general health visits each year.** However, there were differences in the numbers of routine employment support visits. In each year on average, the Full-Service and Basic-Service participants had more employment support visits than the Usual Services. This result is expected, given that both the Full-Service and Basic-Service groups received IPS SE as a core part of both treatment interventions.

Cost-Benefit Analysis

While net costs outweigh benefits at the close of the SED, strong earnings and benefit growth over time show promise.

Exhibit ES-1. The impact of SED on earnings and fringe benefits increased steadily each year



- Participant earnings and fringe benefits increased steadily each year for both Full- and Basic-Service (see Exhibit ES-1). Benefits increased sharply over time. These trends and persuasive evidence from other studies suggest that these 3-year net benefit comparisons grossly understate net benefits over a somewhat longer timeframe.

- Across all sites combined, the average per-participant cost for Full-Service participants was $23,521, and the average Basic-Service cost was $14,483. The main cost driver of both treatments was labor costs ($21,949 and $13,110, respectively), with clinical technical assistance and other behavioral health and work-related expenses making up the rest. The average site-specific cost for Full-Service participants ranged from $14,760 to $28,781 per participant; for Basic-Service participants, the average per-participant costs ranged from $9,541 to $17,087.

- Full-Service and Basic-Service participants received an average of $4,880 and $3,493 in earnings and fringe benefits more than the Usual Services group. Earnings made up about half of the benefits, while fringe benefits from work (health insurance, retirement, vacation and sick time benefits) made up the rest.

- There were no statistically significant cost savings in healthcare utilization costs during the study period. On average, Full-Service participants had $10,277 in combined hospital inpatient stays, ER visits, and outpatient visits during the 3-year study period, compared to $9,378 for Basic-Service and $10,873 for Usual Services participants; the differences between the groups were not significant.

- Over the 3-year study period, the total costs outweighed benefits for both the Full-Service and Basic-Service participants. We estimated a range of net cost-benefits based on different assumptions about costs for the Usual Services group. For the low estimate for the average Full-Service participant, the total costs outweighed benefits by $16,367, and for the average Basic-Service participant, total costs outweighed estimated benefits by $8,714. For the high estimate, we use the upper bound estimate of employment support costs for the Usual Services group ($5,942). For the high estimate for the average Full-Service participant, the total costs outweighed benefits by $12,836, and for the average Basic-Service participant, total costs outweighed estimated benefits by $5,183.

1. Study Overview

1.1 Background

The SSA oversees two programs providing cash benefits to persons with a disability, as determined by the agency: Social Security Disability Insurance (SSDI) and Supplemental Security Income (SSI). The SSDI program provides benefits to disabled workers and their families. In 2017, SSA paid more than $149 billion in SSDI benefits to almost 10.1 million people (Social Security Office, 2017). The SSI program guarantees a minimum level of income support to financially needy individuals who are aged, blind, or disabled. In the same year, SSA paid nearly $55 billion in Federal SSI benefits to about 8.2 million people. Given the large number of individuals who rely on these disability programs to make ends meet and the interest in support for employment efforts, policymakers need an evidentiary base from which to consider potential program improvements and innovations that can strengthen the ability of individuals with disabilities to work.

Persons with a disability due to mental impairment constitute a large proportion of individuals on the disability rolls and represent a major area of policy interest for SSA. The agency previously studied individuals receiving Social Security disability benefits in the Mental Health Treatment Study (MHTS); for these individuals, the MHTS showed that persons with mental impairments can engage in competitive work and that "services can matter" (Frey et al., 2011). Provision of services had positive effects on employment, mental health, and lower hospital utilization. Having detected positive effects for disability beneficiaries, there is a strong interest in applying the same intervention model earlier in the disablement process, with the goal of changing its endpoint.

One potential target of opportunity for increasing participation in the labor force by individuals with a disability is to intervene earlier in the disablement process. Presumably, individuals who apply for disability benefits but receive a denial do not have an impairment that is sufficiently determinable or severe to meet the statutory definition of disability. Across all levels of appeal, SSA denies benefits to about half of those who apply, and a large portion of those denied allege a mental impairment. Statistics indicate that, over time, if these individuals continue to have health problems and labor market problems, several years later they will subsequently reapply and receive approval for benefits associated with either the SSI or SSDI programs. Early statistics for this population with mental impairments reveal that among individuals who were awarded benefits in a given year, nearly one-third of them were re-applicants (Riley et al., 2021).

Providers of services to people with mental impairments may consider them "hard to serve." However, many forms of mental impairments are treatable, and there are promising findings from research on interventions that integrate treatment with vocational services. There is evidence that many applicants with mental impairments do not receive the integrated vocational and behavioral health services that could help them recover and enjoy a more productive life (Jans, Stoddard & Kraus, 2004).

1.2 Design of the SED

The SED is a multicomponent intervention aimed at improving the employment outcomes of Social Security disability applicants alleging a mental impairment who were recently denied benefits. The primary question that SSA seeks to answer is whether offering the IPS model of employment services with behavioral health and other services fosters employment and clinical recovery that leads to self-sufficiency, improved quality of life, and less demand for disability benefits. Findings from this study can help policymakers improve existing programs that may lead to increased participation of individuals with disabilities in the workforce.

The SED used an RCT experimental design with three study arms to compare two treatment alternatives, entitled “Full-Service” and “Basic-Service,” against a “Usual Services” control group as the counterfactual. The SED randomized 3,000 invited participants, aged 18 to 49 years, to one of the three study arms with the goal of 1,000 participants per arm. Participants randomized to one of the two treatment arms receive intervention services from one of 30 organizations serving as demonstration sites across 20 states. The duration of participation in the study was 3 years. Study participants live within the catchment area (i.e., geographical service area) of one of the 30 organizations and all expressed a desire to work or, if employed, sustain or seek a better job. The study excludes individuals who do not have the ability to provide informed consent, those who were already receiving employment services from the demonstration site, and those who resided in a nursing home or other custodial institution.

Exhibit 1-1 summarizes the services available to participants in each study arm. The central feature of the intervention model for the Full-Service and Basic-Service treatment groups is the integration of SE with behavioral health treatment, following the evidence-based IPS model of employment services. IPS has repeatedly demonstrated effectiveness in assisting individuals with severe mental illness to secure and maintain competitive jobs (Bond, Drake & Becker, 2012; Brinchmann et al. 2019; Frederick & VanderWeele, 2019; Kinoshita et al., 2013; Marshall et al., 2014; Metcalf, Drake & Bond, 2018; Modini et al., 2016; Suijkerbuijk et al., 2017). The intervention model for both treatment groups also includes the provision of care management services (e.g., mental health case management, short-term supported education, social skills training, housing assistance, financial assistance, legal assistance, substance abuse counseling and treatment, family counseling, and benefits planning) to address barriers to employment.

Exhibit 1-1. Levels of service

	Full-Service	Basic-Service	Usual Services
Receive manual of available mental health and employment services in their community	✓	✓	✓
Healthcare coverage for those uninsured	✓	✓	✓
Care management services	✓	✓	
Financial support for work and health expenses.	✓	✓	
Medication Management Support	✓		

With an employment-first focus, the Basic-Service and Full-Service treatment groups integrate behavioral health and care management services to eliminate barriers to employment. Both interventions also include modest financial support for individual work-related expenses and out-of-pocket expenses associated with behavioral health and other care management services not covered by health insurance. The sole difference between the two treatments is the availability of services from an NCC for participants in the Full-Service treatment group. Hence, providing NCC services increases the cost of the Full-Service intervention. Because the SED incorporates an RCT design, the evaluation can measure whether this added service enhances participant outcomes and at what cost.

Participants in the Usual Services group seek employment or mental health services on their own, as they would normally in their community. At the time of randomization, each Usual Services enrollee received a comprehensive manual describing mental health and employment services in their local community as well as state and national resources available to them. Regardless of their randomization assignment, all SED participants were eligible to receive healthcare paid by the study if they did not have health insurance at the time of enrollment, or if they lost their health insurance during the study. These payments continued until they could apply for health insurance through a state or federal marketplace by the end of the next Open Enrollment period for the ACA.

1.3 Overview of the Report

This report provides the overall findings on the impact and cost-benefit analysis (CBA) for the SED. We report impacts on outcomes for all study arms, provide net benefits estimation, and discuss challenges, opportunities for further analysis, and overall policy implications.

Including this Introduction (Chapter 1), the report contains seven chapters:

- **Chapter 2 – SED Implementation.** This chapter explores the implementation design for the study, including site selection for the interventions, the target population of denied disability applicants, enrollment and randomization equivalence across the three study groups, and prominent characteristics of the enrolled study population.

- **Chapter 3 – Participation in Research.** In this chapter, we assess the overall rates of participation in the quarterly follow-up surveys and the impact of nonresponse on the outcomes analysis. We describe methods to address nonresponse in the final analysis.

- **Chapter 4 – Impact Evaluation.** This chapter provides an assessment of the effectiveness of the interventions by comparing the outcomes of interest across the Full-Service, Basic-Service, and Usual Services study conditions. The Usual Services arm served as the control group for both the IPS and integrated mental health and related services offered in the Full-Service and Basic-Service treatments.

- **Chapter 5 – Cost-Benefit Analysis.** This chapter presents findings of the SED CBA. We discuss the accounting framework that was used; procedures to monetize benefits and costs; the steps necessary to determine the present value of net benefits; and plans for extrapolating these estimates to future years.

- **Chapter 6 – Synthesis of Process Evaluation and Impact Findings on Outcomes.** This section presents a synthesis of the key process analysis and impact analysis findings, focusing particularly on the research questions originally posed by SSA for the SED.

- **Chapter 7 – Study Limitations, Key Impact Findings, and Policy Implications.** This chapter revisits the SSA's original research questions, identifies study limitations, presents a summary of the key findings, and consolidates the findings and information from the previous chapters into a set of policy implications.

2. SED Implementation

The impetus for the SED was the belief that a multicomponent intervention based around SE could assist denied disability applicants to return to the workforce. The demonstration would offer SSA and the larger society evidence of a strategy to assist working-age people who struggle with declining health to continue their working life and to reduce their immediate demand for disability benefits. This study attempts to evaluate whether intervening with evidence-based healthcare and vocational support would foster employment and clinical recovery among recently denied disability applicants. SSA targeted recently denied disability applicants who alleged mental illness and were between the ages of 18 and 49 in 30 catchment areas in 20 states around the country. The SED tests the proposition that providing such support would lead to higher levels of employment and reduce immediate demand for disability benefits among this unique population. If successful, many meaningful benefits would accrue to these people and their families as well as to society. SSA, in its evaluation, wanted to know whether targeted evidence-based services would benefit this population and whether such an intervention would be cost effective.

Each year people struggling to maintain work in the face of declining health decide to apply for disability benefits through the two Social Security disability programs. Believing their health condition compromises their ability to work, these individuals decide to apply for disability, often at the urging of a family member, health provider, or another concerned person. In 2018, 35 percent of the new applicants for disability benefits (including SSDI, SSI, and concurrent applicants) received awards at the initial determination level. Following the various levels of appeals for those applicants in 2018 who were denied, another 13 percent also received awards. Thus, in 2018, SSA awarded benefits to nearly half (48%) of its new disability applicants (Social Security Office, 2019).

SSA views the SED as an opportunity to intervene early in the disablement process, before individuals with disabilities fully separate from the labor force. Presumably, denied applicants can work, at least above SGA—as their denial implies. They have not been determined to have a health condition (i.e., an impairment) that is sufficiently severe and/or medically determinable to meet the statutory definition of disability. If their health and work problems continue without intervention, over time they may appeal or later reapply for disability benefits and may eventually receive an award. Effective intervention when their initial application is denied may delay or eliminate the need for an appeal or a re-application in the near future.

In this chapter, we explore the implementation design for the study, including site selection for the interventions, the target population of denied disability applicants, enrollment and randomization equivalence across the three study groups, and prominent characteristics of the enrolled study population.

2.1 Site Selection

Thirty community agencies in cities across all seven SSA regions and in 20 U.S. states comprised the setting for the demonstration. We arrived at this number of sites after weighing several key factors, including the expected number of denied disability applicants that a typical community agency could serve for the study period, the potential number of denied disability applicants in specific agency catchment areas that we could recruit for the study, and generalizability of the study results to broader geographic regions. In addition, we wanted sites to represent both Medicaid expansion and non-expansion states, as access to healthcare could become a mitigating issue for study costs.

To reflect national diversity, we sought to include a balance of urban and rural sites, as well as sites that reflected ethnic and racial diversity. To capture a mix of agencies in urban, suburban, and rural areas, we decided to split the demonstration sites into full and small sites, given the only way to include some rural areas would be to expect fewer study enrollees due to the lower number of denied disability applicants residing in those areas.

In our experience with services demonstrations, existing mental health treatment facilities maintain a dynamic balance between the size of their staff and the size of their caseload. We know from experience that client waitlists exist at many programs. A large influx of new clients would result in a temporary imbalance in staffing that would affect treatment quality. Our experience with the MHTS suggested that when enrollments reached high numbers in some sites, the facility experienced exceptional difficulty adjusting to the influx of clients due largely to staffing shortages and scheduling conflicts. After reviewing enrollments across MHTS demonstration sites, we determined that the best course of action would be to rely on no more than two full-time equivalent (FTE) IPS specialists and one FTE NCC at any given site. Smaller, more rural sites could accommodate one IPS specialist and an NCC with one-half FTE. Given caseload expectations, we estimated that the demonstration sites should engage no more than 80 intervention participants (40 Full-Service and 40 Basic-Service) over the 4 years of the study. Thus, we planned 120 participants in each large site: 40 Full-Service, 40 Basic-Service, and 40 Usual Services. Small sites would enroll half those numbers—that is, 60 participants, including 20 Full-Service enrollees, 20 Basic-Service enrollees, and 20 Usual Services enrollees. Adding more participants to fewer agency workloads would likely reduce overall quality of treatment in the facilities, either for study enrollees or for nonstudy clients. Given the expected number of 3,000 study enrollees, we determined a need for 20 large sites and 10 small sites.

We engaged both state Vocational Rehabilitation and Mental Health agencies to obtain recommendations of high-fidelity IPS sites. Thus, all sites selected shared the commonality that their state office(s) perceived them as operating a solid, high quality IPS program. We began with state agencies participating in the existing IPS Learning Community, a collaborative at the time comprising 22 states and over 250 community agencies using the IPS model of SE. We received over 60 nominations for sites that also expressed interest in participating in the demonstration.

The next step was to determine whether potential site catchment areas had enough denied disability applicants within a recent year to make recruitment feasible. Assuming a take-up rate of 20 percent, we determined that we would need catchment areas with at least 600 denied applicants in large sites and 300 denied applicants in small sites during a single year, given an anticipated enrollment period of 12 months. We defined the catchment area for each site using specifications provided by each nominated site describing the geographical area to which they provide services. We requested from SSA data on denied disability applicants for the two most recent years for the catchment area for each site. Based on the results, we classified each site as potentially a large site or a small site and proceeded to engage the sites in formal negotiations with an eye toward maintaining geographical representation across the United States; including both Medicaid expansion and non-expansion states; and including agencies serving urban, suburban, and rural populations. Lastly, we engaged each potential site in discussions about the specific details of the study and gauged their commitment to its implementation. The final demonstration sites appear in Exhibit 2-1, which is a map of the United States with site locations labeled by SSA region.

Exhibit 2-1. SED site locations

2.2 Target Population

The SED enrolled 3,000 recently denied disability applicants living within the catchment area of one of the 30 community agencies contracted by the study to deliver intervention services. After specifying for SSA the ZIP Codes for each catchment area, SSA sent files to Westat files month during the enrollment period containing the names, Social Security numbers, addresses, and telephone numbers of applicants denied in the previous 2 months. Over the course of the enrollment period of 16 months, SSA sent the unduplicated names and addresses of 73,512 denied applicants. Westat carefully documented the disposition of all 73,512 denied applicants in relation to the study. Complete details of the recruitment and enrollment process and results are available in the Enrollment Analysis Report (Taylor, Salkever, Frey, Rile & Marrow, 2020).

From files containing the original 73,512 denied applicants, 26,505 failed a special programmatic screener, leaving a pool of 47,007 denied applicants eligible to participate in the study. The first step in applying the screener included eliminating those cases on the file with flags that met any of three ineligibility criteria: (1) applicants with a code indicating there was no actual determination—meaning they were denied for a reason other than not having a determinable impairment; (2) the file contained information that indicated the denied applicant had an intellectual impairment (i.e., cognitive deficit); or (3) the file indicated the denied applicant's primary language was a language other than English or Spanish. This step eliminated 15,869 otherwise eligible denied applicants.

In a second step, all remaining names and addresses were sent to LexisNexis for tracing and to augment contact information for cases with missing or incomplete address information. Once the file was updated with information from LexisNexis, it was again reviewed to determine whether new address information was valid for the catchment areas. This step resulted in elimination of another 10,636 cases because the new data indicated the denied applicant was (1) incarcerated; (2) living in a residential mental health treatment facility; or (3) now living outside the community agency catchment boundary. In reality, many of these newly determined ineligibles resulted from community agencies reducing their acceptable catchment boundaries to a smaller catchment area due to inability to provide services across their very large geographic areas originally defined by their agency catchment. These agencies decided they could not easily serve study participants in all of their office locales. With intervention teams typically housed in a single office site, potential enrollees residing in areas outside the actual office catchment area (but still within the agency catchment area) would find it difficult to participate in services. For this reason, several agencies reduced their overall catchment area to an "effective" catchment area, thus reclassifying many otherwise eligible denied applicants now ineligible for the study.

Two other criteria further reduced the effective target population. Forty-nine denied applicants on the SSA files became ineligible as SSA recognized they were already participants in another SSA demonstration.[2] Finally, prior to and throughout the enrollment period, an additional 442 cases became ineligible because the denied applicant began receiving disability benefits before they could be recruited into the study. These individuals either appealed their denial and it was accepted before their enrollment, or SSA discovered they received benefits based on a separate prior application.[3] The programmatic screening criteria winnowed the effective target population to 46,516 denied applicants who were eligible for recruitment into the SED.

Each month that Westat received SSA files containing newly denied disability applicants, we followed the steps described above to determine eligible candidates for the study. Once confirmed eligible, all candidate names and addresses were organized at random into release groups of 25. A total of 21,001 randomly selected eligible candidates for enrollment were released for recruitment until the study enrolled and randomized 3,000 study enrollees. The remaining 25,515 eligible candidates across the 30 demonstration sites were never released for recruitment as they were not needed.

We compared basic demographics and other relevant variables between eligible candidates selected for recruitment (21,001) and those eligible candidates not selected for recruitment (25,515). The two groups were very similar in demographic characteristics (age, gender, education), as well as job history, number of jobs held in the past 15 years, and weekly pay of most recent job, with no significant differences noted. The two groups differed on two variables related to their disability status. One difference concerned the medical screen (step 2) in the disability determination process. A higher percentage in the group selected for recruitment (13.3%) were denied at this step than in the reserve group (13.0%). Denial at this step indicates that SSA determined the applicant's impairment was either not severe enough to receive an approval or did not have a sufficient duration. Also, a higher percentage of the reserve group was denied at step 4

---

[2] SSA allows its applicants and beneficiaries to participate in only one agency demonstration.

[3] Fifty-six current disability beneficiaries were later found to be enrolled in the study.

(2.3%) compared to the group selected for recruitment (1.9%). Receiving a denial at step 4 indicates that SSA determined the applicant was capable of working a previous job. The other variable revealing a difference between the two groups of eligible candidates was previous denials for another application. A higher percentage in the group not selected for recruitment (15.4%) received a denial previously for the same type of claim compared to the group selected for recruitment (11.2%). It is not clear what impact either of these two significant variables might have had on either recruitment into the study or the final study outcomes. However, it is worth noting that in both cases the overall numbers are relatively small. Details of these analysis are described in Taylor, Salkever, Frey, Riley & Marrow, 2021.

In 2019, SSA denied SSDI and/or SSI benefits to over 2 million applicants (Social Security Office, 2020a). Among those adjudicated for a diagnosis, approximately 12 percent have a mental disorder. As a rough estimate, the number of denied applicants with a mental disorder in a given year who may be eligible for an intervention such as the SED is approximately 250,000 (Social Security Office, 2020a; Social Security Office, 2020b).

2.3 Enrollment Recruitment, Engagement, Consenting, and Randomization

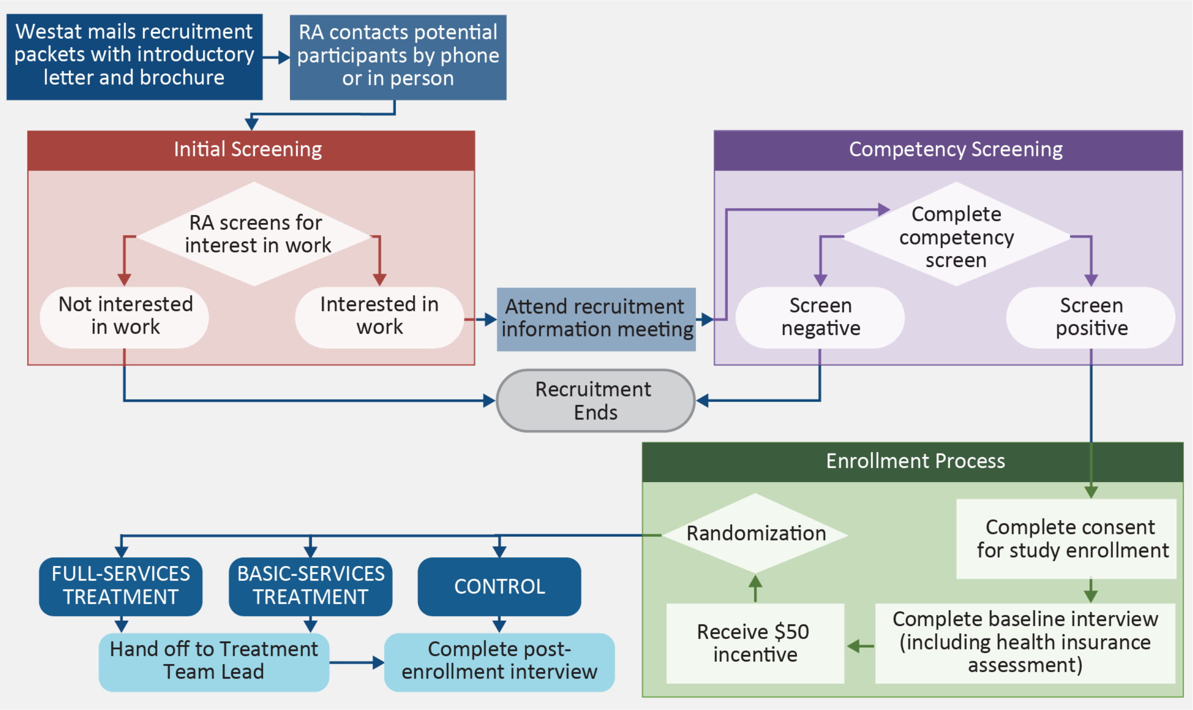
RAs residing in the local community of each community agency conducted all recruitment activities, including promoting the study, engaging, screening, and enrolling eligible denied SSA disability applicants. Westat recruited, hired, and trained 43 RAs as field staff. These individuals would remain independent from the demonstration sites throughout the enrollment period to avoid site contamination with control group study participants. Attempts were made to hire all RAs residing locally, ensuring as much familiarity as possible with their assigned catchment area. The study struggled to hire RAs in two sites and finally had to resort to hiring an RA who was not local. We also hired several additional “traveling” RAs to fill in when needed across the study catchment areas. When sites became overwhelmed with increased study interest or lagged behind enrollment projections, a traveling RA was sent in to assist the existing RA. Some sites required a bilingual (Spanish-speaking) RA. During the RA hiring process, the demonstration sites provided feedback about their community and clientele and suggested either hiring a bilingual RA or providing bilingual support when needed.

Three field supervisors each worked with RAs in 10 demonstration site catchment areas. Supervisors were experienced research staff known to Westat from previous work on major research studies. All were highly accomplished middle managers living in time zones close to their portfolio of sites. The supervisors received training specific to the study along with the RAs. The supervisors monitored RA progress, listened in on meetings and enrollment sessions, and provided assistance, encouragement, feedback, and support to the RAs.

Before beginning work, RAs and supervisors completed a rigorous training and evaluation program to ensure consistency in recruitment across all sites. The RAs received the same training and materials, including an agenda, recruitment and enrollment manual, and copies of the materials needed to recruit potential enrollees. During training, RAs received detailed discussion guides and written informational responses to anticipated questions to supplement the training manuals. The study also provided written informational responses anticipating some enrollee questions, which helped to ensure consistency. Following training and based on feedback received from the field supervisors about the recruitment information meetings, the central office sent RAs updates for their training manuals and scripts as needed throughout the enrollment period, further ensuring dissemination of consistent information to potential study enrollees.

Recruitment was a centralized process with a goal to maintain procedural consistency and improve tracking capabilities. Exhibit 2-2 gives an overview of the steps of the recruitment and enrollment process once the release groups were loaded into the study's Management Information System (MIS) and the RA given access to their specific list. RAs received access only to the recruitment portion of the MIS and only to recruitment data pertaining to their site.

Exhibit 2-2. Participant recruitment, engagement, and enrollment process



Westat mailed initial recruitment materials to potential enrollees on the RA's current list. The mailing included a study introduction letter and recruitment brochure. As the RA worked individual cases, they documented the recruitment status of each study-eligible denied applicant on their current list by adding information to the MIS for each contact attempt. The first contact between the RA and a potential enrollee was by telephone 3-5 days after Westat mailed the initial study invitation package. The RA asked potential enrollees about their interest in work, as this is an important criterion for IPS. Those who expressed a desire to work or to improve their work situation were invited to a recruitment information meeting, where they would present in-depth information about the demonstration. This training gave potential enrollees the opportunity to weigh the merits of the demonstration while allowing the RA to dispel any misconceptions and provide detailed information on the demonstration's advantages and risks. The group (sometimes individual) meetings took place throughout the enrollment period as frequently as needed. If an RA was unable to reach a potential enrollee on the phone, they followed up with an in-person visit to the address on file. As part of the initial contact, the RA screened for initial eligibility to determine interest in work or improving the work situation.

The recruitment strategy included steps to minimize the amount of time between the SSA denial decision on benefit application and enrollment into the study. We expected some potential enrollees to have already appealed their SSA decision before learning about the SED. However, we elected to offer the study to eligible individuals regardless of their appeal status. While we expected a small number of people who expressed a desire to work may have already appealed, we believed including them can provide important information on their experiences and outcomes when faced with a decision to participate in an employment support program. Enrollees in both treatments groups and the control group could appeal SSA's decision at any time during the demonstration. While possible, it was expected to be highly unlikely that a recruit would be receiving disability benefits. If an enrollee volunteered that they were already receiving disability benefits from SSA, they were made ineligible for the study. In reality, SSA discovered early in the study that there were a number of "eligible" supposedly denied disability applicants who were already receiving disability benefits—some of them already enrolled in the study.

Potential enrollees committing to participate in the study for 3 years signed a written consent indicating they understood the risks and benefit of participation. Following the consent process, the RA conducted a baseline interview, which included a health insurance assessment to identify uninsured enrollees and an extensive vocational and health history. Within 2 days of the interview, a member of Westat's central Operations Team contacted those enrollees (in all three study arms) who did not have health insurance to provide them with information on how the study would pay for needed medical care until they were eligible to obtain coverage through the ACA.

Enrollees received modest payments for their time and effort to complete the enrollment process ($50) and subsequently when they completed quarterly ($25) and annual ($40) interviews during their 3 years of study participation.

2.3.1 Randomization Method

Key considerations for selecting a randomization scheme included level of predictability and outcome balance across the three study arms: Full-Service, Basic-Service, and Usual Services groups. **Predictability** in the context of the SED concerned the ability of the RA to "guess" what study arm assignment the next enrollee would receive. We were worried that high predictability would potentially allow the RA to manipulate the assignment process with some enrollees. Thus, low predictability was clearly a desired characteristic of the random assignment method we chose. At the same time, we were concerned about balance among the three study arms. The randomization yields a balanced allocation with assignment of the same number of enrollees (or nearly the same) to each study arm at each site. In addition, we sought to achieve the same number of enrollees in each study arm for four subgroups of interest: including SSI applicants age 18-34; SSI applicants age 35-49; SSDI applicants age 18-34; and SSDI applicants age 35-49. We often view predictability and balance as antagonists in the randomization process, with lower predictability achieved at the cost of reduced balance. Our goal was to employ a method with low predictability and high balance.

In an effort to select a randomization method that would achieve our goal, we conducted a simulation of potential methods using 20 large study sites (120 enrollees per site) and 10 small sites (60 enrollees per site). We tested four methods. The simulation study and theoretical considerations both pointed to the dynamic allocation method as giving the lowest predictability and highest balance at the end of recruitment, and we decided to use that method for randomly assigning enrollees to study arms.

The enrollment process included getting written consent from the prospective study participant and completion of an extensive baseline questionnaire. This process required 1 hour or more. After completing the enrollment, the new study enrollee received an assignment to one of the two treatment groups (Full-Service or Basic-Service) or to the control (Usual Services) group via the randomization program RAs accessed within the MIS.

Handoffs by the RA to the appropriate intervention team was handled in a variety of ways depending upon the comfort of the enrollee. Most often, the RA contacted the relevant intervention team lead while the new enrollee was still in the office to set up an appointment for the new enrollee. However, due to schedules and individual circumstances, that was not always the way contact was made. Sometimes the RA simply gave the enrollee the contact information of demonstration site team lead, as well as sending an email with new enrollee contact information to the team lead and having the site make the contact with the new enrollee.

New enrollees assigned to the Usual Services group were provided a comprehensive manual of mental health and vocational services and support groups in the community and state. The manual also contained the contact information for national support groups.

2.3.2 Enrollment Results

The enrollment period lasted 16 months from the fourth week of November 2017 to the fourth week of March 2019. The original projection was that an enrollment period of 12 months would be needed to achieve 3,000 study participants, with 120 enrollments projected for large sites, and 60 projected for small sites. As the planned enrollment period of 12 months ended, some sites struggled to complete their enrollment quotas within that timeframe. A few sites completed enrollment within the 12-month period, but many did not, requiring a month or more to finish enrolling the allotted number of participants. Among those failing to reach their anticipated enrollment projections were six sites that struggled to reach their expected enrollment. Some ran out of sample, others simply struggled to enroll denied applicants, especially in large rural areas that required substantial travel to interview prospective study participants. Our projections suggested that either working through the existing sample or adding enough new sample given historical enrollment rates were not cost effective for the study. For other sites, RA staffing problems made it easier to transfer the additional sample to sites that were better prepared to enroll additional study participants. Seven sites agreed to use their existing sample to enroll additional study participants. Overall, we diverted 71 enrollment slots from six sites (ranging from 2 to 25 slots at each site) and added them to seven other sites (ranging from 2 to 15 slots).

The enrollment rates ranged from 23.0 percent to 26.2 percent. The lower end of 23 percent included the 3,000 denied applicants who enrolled in the study within the context of the larger population of denied applicants who were eligible for the study and also released for contact. The higher end of 26.2 percent includes the 3,000 enrollments within the context of only those eligible denied applicants for whom we have confirmation that they at least heard about the study and knew they were eligible. For the larger population, we sent letters and attempted to make personal contact but have no data to suggest that they ever personally became aware of the study and their eligibility. They may have received the mailed materials, but we had no way of knowing that.

It is noteworthy that recruiters experienced varying levels of success in recruiting at different sites. At some sites recruiters were unable to speak with 20 percent or more of the eligible denied applicants. At these sites, a larger portion of applicants did not return phone calls and did not make themselves available for the initial screener. Although we cannot know whether the applicant knew about the study, at least a portion of these "possible potential enrollees" received voice mails and other contact attempts and decided not to return the call. Further, among those contacted, the initial screener removed almost half (44.1%) of the eligible denied applicants, mainly because they refused to participate, with no interest in working or finding a new job as the most common reason.

Early in the study implementation, field staff at the intervention sites reported that one or more of their enrollees reported receiving disability benefits. Staff immediately let Westat know about the situation and Westat in turn informed SSA. SSA began investigating the identified cases and found that indeed some enrollees were receiving benefits from another previous application. These applicants denied disability on their "current" application were deemed ineligible for the study. A comprehensive review of the entire study sample released to Westat revealed approximately 160 persons in this category. Among this group were 56 denied applicants already enrolled in the study. While the study continued to serve these enrollees, they were removed from all subsequent analyses since they were technically ineligible. These events reduced the effective sample size from 3,000 to 2,944.

2.4 Characteristics of Enrolled Study Participants

Table 2-1 provides data on the characteristics of the eligible SED enrollees. The average age of all enrollees was 36.1 years with 57.6 percent reportedly between the ages of 35 and 49. The majority of enrollees were White, non-Hispanic (48.3%) with 28.3 percent reporting as Black, non-Hispanic, and another 12.4 percent reporting as Hispanic. Over half (56.5%) of the eligible enrolled study population were female. A majority of enrollees (51.2%) reported some college, an associate, bachelor's, or higher degree. Another 30.2 percent attained a high school diploma or general education diploma (GED), and 18.6 percent reported education as having not received a high school diploma. Nearly two-thirds of enrollees' applications (64.4%) reflected work histories meriting the SSDI claim type. The remaining 35.6 percent of enrollees' claim type was for the SSI program. This latter group likely fits into the category of the 35.8 percent of enrollees who reported never working or they hadn't worked in the 2 years prior to enrollment in the study.

Analysis of baseline data for the 2,944 eligible enrolled study participants also revealed an unstable study population with precarious life situations. As anticipated, the self-reported mental health status of the overall study population was poor (32.6), falling nearly 2 standard deviations below the average of 50 on the SF-12 Mental Component Score (MCS). However, very surprising was the poor physical health of enrollees. The average physical component score on the SF-12 was 38.3 reflecting a score of more than one standard deviation below the mean of 50. Poor health may also be a consequence of the fact that 21 percent of enrollees reported they had no health insurance. In addition to poor health and limited access to healthcare, the living situations reported by enrollees present unusual challenges. Over 80 percent of enrollees reported having never been married or were currently separated, divorced or widowed. The same statistic for the overall U.S. population was only 59.9 percent, more than 20 percent lower than the study population. In addition, enrollees reported a poverty-level average household income of $1,847 in the month before enrollment—more than 70 percent lower than the average U.S. household income of $6,774.

Table 2-1. Characteristics of eligible SED enrollees by study arm

Variable	Full-Service (976)		Basic-Service (987)		Usual Services (981)		Full sample (2,944)		p-value
Age (years)	n	M	n	M	n	M	N	M	0.92
Mean	976	36.14	987	35.97	981	36.04	2,944	36.05	
Median	976	37	987	37	981	37	2,944	37	
Age group	n	%	n	%	n	%	n	%	0.95
18-34	411	42.11	422	42.76	414	42.20	1,247	42.36	
35 and above	565	57.89	565	57.24	567	57.80	1,697	57.64	
Claim type	n	%	n	%	n	%	n	%	0.86
SSDI	632	64.75	629	63.73	635	64.73	1,896	64.40	
SSI	344	35.25	358	36.27	346	35.27	1,048	35.60	
Gender	n	%	n	%	n	%	N	%	0.21
Male	405	41.50	429	43.47	446	45.46	1,280	43.48	
Female	571	58.50	558	56.53	535	54.54	1,664	56.52	
Race and ethnicity									0.01
White non-Hispanic	503	51.54	440	44.58	478	48.73	48.27	1,421	
Black non-Hispanic	243	24.90	314	31.81	277	28.24	834	28.33	
Hispanic	114	11.68	137	13.88	113	11.52	364	12.36	
Two or more races non-Hispanic	91	9.32	72	7.29	83	8.46	246	8.36	
Other or missing	25	2.56	24	2.43	30	3.06	79	2.68	
Education	n	%	n	%	n	%	N	%	0.84
Less than high school	197	20.18	175	17.73	176	17.94	548	18.61	
High school or GED	280	28.69	308	31.21	301	30.68	889	30.20	
Some college	329	33.71	343	34.75	337	34.35	1,009	34.27	
Associate	67	6.86	70	7.09	69	7.03	206	7.00	
Bachelor's or higher	103	10.55	91	9.22	98	9.99	292	9.92	
Access to reliable transport	n	%	n	%	n	%	N	%	
Yes	762	78.07	793	80.34	776	79.10	2,331	79.18	0.78
No	205	21.00	186	18.84	198	20.18	589	20.01	
Refused/don't know	9	0.92	8	0.81	7	0.71	7	0.71	

Table 2-1. Characteristics of eligible SED enrollees by study arm (continued)

Variable	Full-Service (976)		Basic-Service (987)		Usual Services (981)		Full sample (2,944)		p-value
Work history	n	%	n	%	n	%	N	%	0.36
Working at baseline	173	17.73	199	20.16	189	19.27	561	19.06	
Worked in past 2 years	441	45.18	444	44.98	430	43.83	1,315	44.67	
Worked but not in past 2 years	328	33.61	317	32.12	328	33.44	973	33.05	
Never worked	33	3.38	21	2.13	28	2.85	82	2.79	
Refused/don’t know	1	0.10	6	0.61	6	0.61	13	0.44	
Have health insurance	n	%	n	%	n	%	N	%	0.41
Yes	789	80.84	787	79.74	770	78.49	2,364	79.69	
No	181	18.55	191	19.35	207	21.10	207	21.10	
Refused/don’t know	6	0.61	9	0.91	4	0.41	4	0.41	
Obese	n	%	n	%	n	%	N	%	0.79
Yes	448	45.90	467	47.32	461	46.99	1,376	46.74	
No	516	52.87	513	51.98	510	51.99	1,539	52.28	
Refused/don’t know	12	1.23	7	0.71	10	1.02	29	0.99	
Housing situation									0.33
At one address	849	86.99	856	86.73	845	86.14	2,550	86.62	
More than one address	83	8.50	72	7.29	80	8.15	235	7.98	
Homeless shelter or on street	43	4.41	59	5.98	53	5.40	155	5.26	
Refused/don’t know	1	0.10	0	0	3	0.31	4	0.14	
Marital status									0.32
Never married	517	52.97	551	55.83	537	54.74	1,605	54.52	
Married or living as married	208	21.31	174	17.63	185	18.86	567	19.26	
Separated/divorced/widowed	250	25.61	259	26.24	254	25.89	763	25.92	
Refused/don’t know	1	0.10	3	0.30	5	0.51	9	0.31	
Ever been arrested	n	%	n	%	n	%	N	%	0.54
Yes	516	52.87	534	54.10	521	53.11	1,571	53.36	
No	452	46.31	450	45.59	456	46.48	1,358	46.13	
Refused/don’t know	8	0.82	3	0.30	4	0.41	15	0.51	

Table 2-1. Characteristics of eligible SED enrollees by study arm (continued)

Variable	Full-Service (976)		Basic-Service (987)		Usual Services (981)		Full sample (2,944)		p-value
SF-12 MCS	n	M	n	M	n	M	N	M	0.12
Mean	951	32.04	960	33.22	965	32.55	2,876	32.60	
SF-12 PCS	n	M	n	M	n	M	N	M	0.90
Mean	951	38.12	960	38.33	965	38.37	2,876	38.27	

These study participants bore little resemblance to the participants of the original seminal IPS studies where participants lived relatively stable lives outside their struggles with mental illness. While participants of the seminal IPS studies had diagnosed serious mental illness that was considered relatively stable, those in the SED exhibited less severe but more complex mental problems associated with untreated personality or anxiety disorders.

2.4.1 Random Assignment Results

Random assignment of the original 3,000 study participants resulted in 994 (33.1%) study participants assigned to the Full-Service treatment group; 1,004 (33.5%) assigned to the Basic-Service treatment group; and 1,002 (33.4%) assigned to the Usual Services (control) group. The random assignment model took into account age group (18-34 vs. 35-49) and disability claim type (SSDI vs. SSI). The resulting assignments revealed no differences between the study arms on either of these variables, or on other key variables, including age, age grouping, gender, education, work history, health insurance status, marital status, arrest history, physical or mental health scores on the SF-12, or household income in the past month. One variable, race and ethnicity, did appear as significant across the three study arms. Fewer White participants received assignments to the Basic-Service group, while more Black and more Hispanic participants received assignments to that group compared to either the Full-Service treatment group or the Usual Services (control) group.

When the 56 ineligible enrollees were removed from the analytic dataset, the resulting balance between the three study groups did not effectively change: 976 Full-Service (33.2%); 987 Basic-Service (33.5%); and 981 Usual Services (33.3%). A reanalysis of differences between the three study arms on the same key variables identified above revealed the same results, with no differences between the study arms on all variables except race and ethnicity. The Basic-Service treatment group was overrepresented by Black and Hispanic participants compared to the Full-Service treatment group and the Usual Services (control) group.

2.5 Assessment of Diversity in Enrollment

After recruitment closed, the team conducted an enrollment analysis to identify characteristics associated with the decision to enroll in the SED (Taylor et al., 2020). This analysis identified several characteristics of denied applicants associated with enrollment; for example, males, individuals with less prior work experience or less earnings, and individuals with greater educational attainment were more likely to enroll in the study.

A key question is whether the SED study enrollees represent the racial and ethnic diversity of the target population (i.e., denied applicants with a mental impairment). A program such as the SED should be available and utilized by a diverse representation of denied applicants. However, very little information is available about the demographics of the population of denied applicants. SSA did not capture complete race or ethnicity information during the application process. Because of the lack of complete data, race and ethnicity were not included as potential predictors of enrollment in the enrollment analysis.

Although we do not know the race and ethnicity of all denied applicants, it is possible to use their location to identify the racial and ethnic diversity of the areas where they live using data from the U.S. Census. We took this approach to compare the racial and ethnic composition of the populations for the census tracts of residence of enrolled versus non-enrolled denied disability applicants included in the SED.

Table 2-2 summarizes the average racial and ethnic demographics of the census tracts of denied applicants who enrolled and who did not enroll in the SED. In terms of race, the census tracts of enrollees were similar to those of non-enrollees, although there are some minor differences to note. SED enrollees on average came from areas with slightly larger proportions of Black residents (20.29% for enrollees, 17.71% for non-enrollees) and slightly smaller proportions of White residents (66.62% for enrollees, 68.47% for non-enrollees). These were the largest substantive differences in race and ethnicity; other differences were within 0.6 percentage points.

Table 2-2. Race and ethnicity of census tracts of enrolled and not enrolled denied applicants in SED catchment areas

	Enrolled n = 2,944	Not enrolled n = 18,059	All n = 21,003	p-value
Race				
White	66.62	68.47	68.21	< 0.01
Black	20.29	17.71	18.07	< 0.01
Native American	0.72	0.84	0.82	< 0.01
Asian	4.03	4.20	4.18	0.14
Native Hawaiian	0.13	0.13	0.13	0.50
Other	4.12	4.41	4.37	0.04
Two or more races	4.10	4.24	4.22	0.02
Ethnicity				
Hispanic	14.41	15.01	14.93	0.07

3. Participation in Research

To measure outcomes and impacts of the interventions, the study team asked all participants in the SED to participate in quarterly interviews throughout their enrollment in the study. Each participant completed a baseline survey and had the opportunity to complete up to 12 quarterly interviews (one for each quarter of their 3-year enrollment in the SED). The study team relied on the data from the quarterly interviews to measure key outcomes and assess the impacts of the interventions.

The study team encountered challenges in locating, contacting, and collecting quarterly interviews. The loss of data from survey nonresponse presents a challenge to the study. In this chapter, we assess the overall rates of participation in the quarterly surveys and the impact of nonresponse on the outcomes analysis. We then describe methods to address nonresponse in the final analysis.

3.1 Outcome Data Sources

We measured the impact of the Full-Service and Basic-Service interventions on outcomes in five domains: employment and earnings; SSA benefit receipt; health status; quality of life; and healthcare utilization. The only outcome we measure using external data outside of the participant surveys is SSA benefit receipt. Outcomes in the domains of employment and earnings; health status; quality of life; and healthcare utilization rely on participant response and recall.

Table 3-1 describes the frequency of data collection for data in each outcome domain throughout study enrollment. The team collected employment and earnings data and healthcare utilization data at each quarterly interview. If a participant skipped a prior quarterly interview, the interviewer asked them to fill in information (e.g., employment start and stop dates) going back to the date of the last completed interview. This approach provides for data that are more complete for the entire study period. The annual surveys (collected at quarters 4, 8, and 12) included health status and quality of life measures. The baseline and quarterly surveys appear in Appendix A and Appendix B, respectively.

Table 3-1. Data collection frequencies for outcome domains relying on participant survey response

Outcome domain	Quarter											
	1	2	3	4	5	6	7	8	9	10	11	12
Employment and earnings	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Health status				✓				✓				✓
Quality of life				✓				✓				✓
Healthcare utilization	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

3.2 Attrition/Research Participation in Follow-Up Surveys

The SED study team measured employment, health status, healthcare utilization, and quality of life outcomes using follow-up surveys. Table 3-2 summarizes the quarterly survey response rates for quarters 1-12. The five columns appearing under the label “All” show the number of completes, deceased, skipped, and withdrawn participants for each quarter. The numbers of deceased and withdrawn increased over time; by quarter 12, there were 83 deceased participants and 50 participants who formally withdrew at some time prior to the interview.

Table 3-2 shows the response rates for each quarter excluding deceased and withdrawn participants. Response rates among eligible participants held above 70 percent for the first 2 years of study enrollment (quarters 1-8). The third year of the study saw a drop-off in completion rates; by quarter 12, roughly two-thirds (65.3%) of eligible enrollees completed the survey.

Table 3-2. Response rates by quarterly interview for eligible enrollees

Quarter	All					Excluding deceased and withdrawn		
	Complete	Deceased	Skipped	Withdrawn	Total	Complete	Skipped	Total
1	2,195	4	739	6	2,944	2,195	739	2,934
	(74.56)	(0.14)	(25.1)	(0.20)		(74.81)	(25.19)	
2	2,070	11	856	7	2,944	2,070	856	2,926
	(70.31)	(0.37)	(29.08)	(0.24)		(70.75)	(29.25)	
3	2,071	17	847	9	2,944	2,071	847	2,918
	(70.35)	(0.58)	(28.77)	(0.31)		(70.97)	(29.03)	
4	2,080	27	820	17	2,944	2,080	820	2,900
	(70.65)	(0.92)	(27.85)	(0.58)		(71.72)	(28.28)	
5	2,105	33	786	20	2,944	2,105	786	2,891
	(71.50)	(1.12)	(26.70)	(0.68)		(72.81)	(27.19)	
6	2,094	38	787	25	2,944	2,094	787	2,881
	(71.13)	(1.29)	(26.73)	(0.85)		(72.68)	(27.32)	
7	2,061	50	802	31	2,944	2,061	802	2,863
	(70.01)	(1.70)	(27.24)	(1.05)		(71.99)	(28.01)	
8	2,043	56	807	38	2,944	2,043	807	2,850
	(69.40)	(1.90)	(27.41)	(1.29)		(71.68)	(28.32)	
9	1,959	62	879	44	2,944	1,959	879	2,838
	(66.54)	(2.11)	(29.86)	(1.49)		(69.03)	(30.97)	
10	1,915	67	916	46	2,944	1,915	916	2,831
	(65.05)	(2.28)	(31.11)	(1.56)		(67.64)	(32.36)	
11	1,875	76	945	48	2,944	1,875	945	2,820
	(63.69)	(2.58)	(32.1)	(1.63)		(66.49)	(33.51)	
12	1,835	83	976	50	2,944	1,835	976	2,811
	(62.33)	(2.82)	(33.15)	(1.70)		(65.28)	(34.72)	

Notes: Percentages are in parentheses, summing across each row.

3.3 Developing Empirical Definitions of Attrition and Analysis of Attrition Rates

We measure four different types of outcomes: (1) outcomes in the domains of employment and earnings, (2) outcomes relating to healthcare utilization, (3) outcomes relating to health status and quality of life, and (4) SSA benefit outcomes (approvals, appeals, and benefits paid). For the first two types of outcomes, we use data from the 12 quarterly follow-up surveys. For outcomes relating to health status and quality of life, we use responses to additional questions included in the quarterly interviews for the end of each full year of the intervention (i.e., quarters 4, 8, and 12). For SSA benefit outcomes, SSA administrative records provide complete data for all SED participants. The follow-up surveys, however, provide data for only those study participants who completed the surveys.

3.3.1 Accommodating Missed Quarterly Interviews

The quarterly follow-up surveys allowed measure of the employment-related and healthcare utilization outcomes over the entire intervention period. The study team expected, however, that many participants would miss at least some follow-up interviews throughout the 3-year study period. Therefore, interviewers asked participants to fill in gaps in employment-related and health services outcomes due to missed prior surveys. For example, if a respondent missed the previous quarterly survey, on the next completed survey the interviewer asks the participant to provide work history for the entire 6 months since the last completed survey (not just the current quarter).

Although the survey design reduced missing data due to skipped surveys by asking respondents to fill in information for skipped surveys, there are two limitations to note. First, participants who broke off and never completed the final survey before transitioning off the study have missing data for all quarters after the final completed survey. Second, participants who missed a large number of surveys must provide information such as dates of employment, hospitalizations, and ER visits for an extended period of time to fill in the gap. A lengthy recall period may lead to inaccurate estimates.

To address these limitations, the study team limited the analysis of outcomes to those participants who completed enough surveys to provide an accurate accounting of their work and health over the course of the study. The outcomes that rely on respondent recall are employment, weeks worked, hours worked, total earnings, hospitalizations, and ER visits. For these outcomes, the analysis includes participants who meet the following criteria:

1. Participants must complete the final (quarter 12) follow-up survey, and
2. Participants must not have a gap in completed quarterly surveys of longer than 1 year.

The rationale behind the second criterion is that research shows that recall periods of less than 1 year can provide reliable estimates of aggregate numbers of events such as hospitalizations (see, for example, Kjellsson, Clark & Gerdtham, 2014).

The team adopted a different approach for the health status and quality of life data reported in the quarter 4, 8, and 12 interviews. For these measures, relevant questions asked for single-point-in-time responses; no gaps due to missing any of these interviews could be filled in during other quarterly interviews (since they did not include the questions relevant to these outcomes).

Thus, health status and quality of life outcomes were measured as the changes between scores measured at baseline and at the quarter 4, 8 and 12 follow-ups. Accordingly, these outcomes include only responses from all participants who completed the relevant follow-up survey items from those specific quarterly interviews.

Completion Rates for Outcome Analyses

Table 3-3 shows completion rates by study year and for the group that finished the final survey and had no gaps in surveys longer than 1 year.

Table 3-3. Completion by year and recall by study arm

	Completed Q4 N = 2,080	Completed Q8 N = 2,043	Completed Q12 N = 1,835	Completed Q12 and no survey gaps over 1 year; N = 1,722
All eligible participants				
Percent of non-deceased	71.31	70.74	64.14	60.19
Study arm				
Basic-Service	73.95	72.89	66.98	62.59
Full-Service	71.64	70.52	65.27	60.88
Usual Services	68.31	68.79	60.13	57.07

Notes: Percentages exclude participants who were deceased at the time of the survey from the denominator.

For measuring outcomes based on recall such as employment, earnings, and healthcare utilization, 60.19 percent of enrollees completed the final survey and had no survey gaps over 1 year. Larger percentages completed the surveys needed to measure health status and quality of life annually throughout the study: 71.31, 70.74, and 64.14 percent completed the year 1, year 2, and year 3 surveys, respectively.

The Basic-Service and Full-Service study participants were more likely to complete the surveys than the Usual Services participants. In each year, the Basic-Service study participants had the highest completion rates, followed by the Full-Service and then the Usual Services participants. The differences in completion rates among the study arms are present in each year of the study but are larger in the final year than in the first year.

Table 3-4 provides a summary of the potential sample size for each analysis by outcome given the response rates among all study participants. Outcomes that rely on recall of dates throughout study enrollment (e.g., employment, healthcare utilization) use data from participants who completed the final interview and had no gaps in surveys over 1 year (n = 1,722). Outcomes measured at specific points in time, such as MCS, include all participants who provided those scores on the annual interviews.

Table 3-4. Outcome measures, definitions, and data sources

Impact measure	Definition	Data source	Sample
Employment			
Employment rate	Percentage of participants who worked during study enrollment	Participant interviews	Completed Q12 survey with no survey gaps over 1 year N = 1,722
Weeks employed	Number of weeks of participant-reported employment	Participant interviews	
Total hours worked	Sum of reported hours worked across all jobs	Participant interviews	
Total earnings	Sum of earnings based on reported work and pay	Participant interviews	
Earnings in the past month	Sum of earnings reported in the most recent month	Participant interviews	
SSA Disability Benefits			
Allowance rate	Percentage of participants accepted onto the disability rolls	SSA administrative data	All eligible enrollees N = 2,944
Time to award	Number of days to disability award	SSA administrative data	
Benefit appeal attempts	Number of appeal attempts during study	SSA administrative data	
Total disability payments	Total disability award payments during study period	SSA administrative data	
Health Status			
Mental Health (SF-12)	MCF difference scores (Study Entry vs. Exit Interview)	Participant interviews	Completed Q12 survey N = 1,835
Physical Health (SF-12)	PCF difference scores (Study Entry vs. Exit Interview)	Participant interviews	
Colorado Symptom Index	Difference scores (Study Entry vs. Exit Interview)	Participant interviews	
Quality of Life			
Satisfaction with life	Difference scores (Study Entry vs. Exit Interview)	Participant interviews	Completed Q12 survey N = 1,835
Utilization of Services			
ER visits	Total ER visits during study	Participant interviews	Completed Q12 survey with no survey gaps over 1 year N = 1,722
Hospital overnight stays	Total overnight hospital stays during study	Participant interviews	
Outpatient visits	Total outpatient visits during study	Participant interviews	
Routine mental health visits	Total mental health visits in last month of study enrollment	Participant interviews	Completed Q12 survey N = 1,835
Routine general health visits	Total general health visits in last month of study enrollment	Participant interviews	
Routine employment support visits	Total employment support visits in last month of study enrollment	Participant interviews	

3.4 Descriptive Data on Participation and Nonparticipation in Follow-Up Survey

We followed three approaches to test and adjust for attrition patterns when estimating impacts: nonresponse weighting, regression adjustment, and bounding. First, we conducted exploratory regressions to identify characteristics associated with attrition and simultaneously control for multiple characteristics that may relate to attrition. We began with baseline characteristics collected from all study participants and identified characteristics related to completion of the final survey and completion of the necessary surveys to construct recall measures. These regressions identified several variables that the team used in the weighting and regression analyses. Variables related to attrition included:

- Gender,
- Age,
- Temporary Assistance for Needy Families (TANF) receipt,
- Supplemental Nutrition Assistance Program (SNAP) receipt,
- Time spent in shelter or on street prior to enrollment,
- Work history (e.g., working at baseline, worked in the past 2 years prior to enrollment),
- Physical Component Score (PCS) baseline score,
- Number of outpatient visits in year prior to enrollment,
- Arrests prior to enrollment, and
- Race and ethnicity.

Table 3-5 summarizes characteristics of enrollees by subgroups based on completion rates. Generally, women, older participants, those with higher levels of formal education, and Black non-Hispanic participants had higher response rates than others. Participants with lower PCS had higher response rates than those with higher PCS.

3.5 Weighting Adjustments

We created nonresponse adjusted weights to estimate outcomes for each study arm. As described above, only participants who responded to more surveys were included in measures that required recall over the entire study period. These recall outcomes required a response to the final (quarter 12) interview and no gaps in surveys over 1 year throughout the 3-year study period. Outcomes that we measured at a single point in time, such as health status at the end of the study, required only the completion of the final (quarter 12) interview.

Table 3-5. Completion rates by subgroup among non-deceased SED enrollees

	Completed Q4		Completed Q8		Completed Q12		Completed Q12 and no survey gaps over 1 year	
	n	%	n	%	n	%	n	%
Gender								
Male	850	67.03	828	66.13	738	59.23	692	55.54
Female	1,230	74.59	1,215	74.27	1,097	67.93	1,030	63.78
Age								
18-34	834	67.42	816	66.07	58.96	58.96	673	54.80
35 and over	1,246	74.17	1,227	74.23	68.03	68.03	1,049	64.24
Education								
Less than high school	354	65.68	349	65.48	325	61.67	292	55.41
Completed high school	611	69.04	619	70.42	539	61.88	510	58.55
Some college	754	75.40	725	73.16	652	66.33	619	62.97
Associate	155	75.98	149	73.76	132	66.33	122	61.31
Bachelor's or higher	206	71.28	201	71.02	187	66.55	179	63.70
Work history at baseline								
Working when enrolled	390	70.02	391	70.32	361	65.05	336	60.54
Not working when enrolled but worked in past 2 years	917	70.38	901	69.84	800	62.60	753	58.92
Worked but not in the past 2 years	710	73.65	686	72.21	614	65.60	578	61.75
Never worked	57	71.25	59	74.68	54	68.35	51	64.56
Missing	6	46.15	6	46.15	6	46.15	4	30.77
Race and ethnicity								
White non-Hispanic	970	68.89	958	68.82	856	61.98	794	57.49
Black non-Hispanic	624	75.54	600	73.35	545	67.37	513	63.41
Hispanic	247	68.99	253	71.07	230	65.53	219	62.39
Two or more races non-Hispanic	183	74.39	181	74.18	154	63.37	148	60.91
Other/missing	56	72.41	51	63.79	50	61.40	48	59.65
Benefit Receipt								
Received SNAP	1,460	72.85	1,444	72.75	1,286	65.55	1,213	61.82
Received TANF	214	70.39	210	69.77	181	60.74	171	57.38

Table 3-5. Completion rates by subgroup among non-deceased SED enrollees (continued)

	Completed Q4		Completed Q8		Completed Q12		Completed Q12 and no survey gaps over 1 year	
	n	%	n	%	n	%	n	%
Mental Component Score (MCS)								
Less than 20	353	68.28	349	68.16	322	63.14	296	58.04
21 to 30	659	72.90	639	71.48	573	64.67	542	61.17
31 to 40	522	73.83	510	72.86	445	64.68	425	61.77
41 to 50	316	69.45	318	70.35	297	66.29	276	61.61
51 or greater	184	69.17	182	68.68	164	61.89	151	56.98
Missing	46	67.65	45	69.23	34	53.13	32	50.00
Physical Component Score (PCS)								
Less than 20	159	74.65	155	73.11	144	68.25	135	63.98
21 to 30	560	74.47	540	72.87	494	67.67	462	63.29
31 to 40	532	72.78	550	75.76	486	67.59	457	63.56
41 to 50	412	71.65	401	69.98	363	63.91	336	59.15
51 or greater	371	64.19	352	61.65	314	55.18	300	52.72
Missing	46	67.65	45	69.23	34	53.13	32	50.00
Criminal Justice Involvement								
Arrested in past year	247	64.32	245	65.16	214	57.53	198	53.22

Notes: Percentages exclude from the denominator participants who were deceased at the time of the survey.

We also analyzed outcomes at each intermediate anniversary of study enrollment (Years 1 and 2). To accommodate these analyses, we used four different definitions of nonresponse to calculate four separate weights:

1. **Weight 1.** Respondents include any study participants who responded to the final (quarter 12) survey. Weight 1 is suitable for analyses of outcomes measured at the end of the study period that do not require recall over the entire study period (e.g., MCS, PCS, Colorado Symptom Index [CSI], earnings in the past month).
2. **Weight 2.** Respondents include any study participant who responded to the final (quarter 12) survey and who did not have a gap in interviews longer than 1 year throughout the study period. Weight 2 is suitable for analyses that rely on recall (e.g., employment, total earnings, total hours worked, number of inpatient hospital stays).
3. **Weight 3.** Includes as a respondent any study participant who responded to the first annual (quarter 4) interview.
4. **Weight 4.** Includes as a respondent any study participant who responded to the second annual (quarter 8) interview.

We calculated nonresponse adjusted weights using data from the baseline survey identified from exploratory regressions of attrition listed above. The team used a Chi-squared Automatic Interaction Detection (CHAID) model to identify variables related to nonresponse using each of the four definitions of nonresponse.

All enrolled SED participants received an initial weight of 1.0. The team then applied adjustment factors to increase the initial weights of responding participants upward for the nonresponding participants.

3.6 Regression Adjustment Approach

We used multivariate regressions to make outcomes comparisons between the study arms and calculate impact estimates. This approach allows for precise and unbiased estimates of treatment impact by controlling for variables that (1) potentially predict attrition, and/or (2) explain variation in the outcome measures. By explaining variation in outcomes, the regressions allow for estimates of impact that are more precise than direct comparisons of means. We use regressions to generate estimates for each pairwise comparison between the study arms (e.g., Full-Service vs. Usual Services, Basic-Service vs. Usual Services, and Full-Service vs. Basic-Service).

To select the variables for the regression model, we began with a list of factors measured at baseline and through external data sources that may explain variation in the outcomes. These factors included participant characteristics collected through the baseline survey related to work status and history prior to enrollment, demographic characteristics, and health status. We also brought in external data that the team matched to the home address of each participant. These data include indicators from the Bureau of Labor Statistics (BLS) of the state of the local economy while the participant was in the study. Finally, we tested indicators of COVID-19 severity during study enrollment using data from the Centers for Disease Control and Prevention (CDC) matched to the participant's home county.

To identify variables for the final models, we used a stepwise approach, including all variables in an initial regression, then removing variables with $p>0.20$. We repeated this process for each of the outcomes. If a variable remained after the stepwise regression (e.g., $p < 0.20$) in at least one of the outcomes models, it was included in the final model for all outcomes.

Table 3-6 summarizes all variables that were included in the initial models. Generally, variables that showed the strongest relationship with outcomes related to work and earnings history, age, and health status at the time of study enrollment.

Table 3-6. Regressions model covariates

Variable	Description	In final model
Work history (baseline survey)		
Work_status_bl	1 = Working at baseline 2 = Not working but worked in past 2 years 3 = Worked but not in past 2 years 4 = Never worked	✓
Weeks_employed_q0	Number of weeks employed in past 2 years	✓
Total_hours_worked_q0	Total hours worked in past 2 years	No
Total_earnings_q0	Total earnings in past 2 years	✓
Demographics (baseline survey)		
DM_2	Gender	✓
Age_baseline	Age at baseline (in years)	✓
DM_14	Received food stamps or SNAP in past 12 months	✓
DM_15	Received TANF in past 12 months	✓
DM_17	Days in shelter or on street in past 3 months	No
Nights_in_jail_q0	Nights spent in jail in past 12 months	✓
Race_eth_comb_miss	Race and ethnicity categories: 1 = White non-Hispanic 2 = Black non-Hispanic 3 = Hispanic 4 = Two or more races non-Hispanic 5 = Other/missing	✓
Urban	=1 if lived in urban area, =0 otherwise	✓
RandomizationDate_dt	Randomization date (date enrolled in SED)	No
EA_EDUC	Highest level of education completed 1 = Less than high school 2 = Completed high school 3 = Some college or technical school 4 = Associate degree 5 = Bachelor’s degree or higher	✓

Table 3-6. Regressions model covariates (continued)

Variable	Description	In final model
Health status (baseline survey)		
MCS_q0	SF-12 MCS at baseline	✓
PCS_q0	SF-12 PCS at baseline	✓
HC_16_R_q0	Outpatient procedures in past 12 months	No
BasicMobilityScore_q0	WD-FAB Basic Mobility domain score	✓
UpperBodyFunctionScore_q0	WD-FAB Upper Body Function domain score	✓
UpperExtremityFineMotor_q0	WD-FAB Upper Extremity Fine Motor domain score	No
CommunityMobilityDrive_q0	WD-FAB Community Mobility Drive domain score	✓
CommunityMobilityRide_q0	WD-FAB Community Mobility Ride domain score	No
CommunicationCognition_q0	WD-FAB Communication and Cognition domain score	✓
Resilience_q0	WD-FAB Resilience domain score	✓
InterpersonalInteractions_q0	WD-FAB Interpersonal Interactions domain score	✓
MoodEmotions_q0	WD-FAB Mood and Emotions domain score	✓
External data		
AverageWeeklyWage_Private_q1_q12	BLS QCEW county Average Weekly Wage, averaged over quarter 1 to quarter 12 of study enrollment	✓
UnemploymentRate_q1_q12	BLS Local Area Unemployment Statistics (LAUS) county Unemployment rate, averaged over quarter 1 to quarter 12 of study enrollment	✓
Hospitalized_rate_avg	CDC state hospitalization rate for COVID-19, averaged over the 36 months of participant's study enrollment for state of residence	✓

4. Impact Evaluation

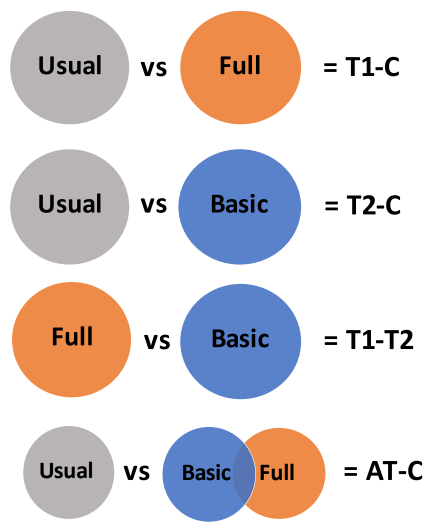
4.1 Introduction

Key Findings

- Full-Service and Basic-Service participants had significantly higher rates of employment than the Usual Services participants during the 3-year study period and in each year of study enrollment.
- The Full-Service and Basic-Service interventions positively impacted total earnings during study enrollment. On average relative to the Usual Services group, about 60 percent of the total SED positive impact on earnings occurred in the third year of study enrollment.
- Employment and earnings impacts differed for certain demographic subgroups. Racial and ethnic minorities saw greater positive impacts from the SED than White participants. Those with lower levels of formal education saw greater employment and earnings impacts from participation in one of the SED treatment groups than those with a college degree.
- The Full-Service and Basic-Service treatment did not have a significant impact on allowance rates, health status, or quality of life.

The impact analysis reveals the extent to which the treatment conditions affect the outcomes of interest following 3 years of intervention. The results allow for an assessment of effectiveness of the interventions by comparing participant performance across the Full-Service, Basic-Service, and Usual Services study conditions. The Usual Services arm served as the control group for both the IPS and integrated mental health and related services offered in the Full-Service and Basic-Service treatments. The Basic-Service study arm offered treatment without the services of a dedicated NCC, allowing for assessment of the impact of having the dedicated nurse on the Full-Service team. A key policy implication of not having the dedicated nurse reduces the cost of the intervention package by about 40 percent. Thus, the comparison between the Full-Service study arm with both the Basic-Service and Usual Services study arms offered an opportunity to assess the added benefit of the NCC, who provided medication management and coordinated overall health services to participants in the Full-Services study arm. Exhibit 4-1 provides a summary of the comparisons.

Exhibit 4-1. Primary comparison groups



The impact analysis includes four primary comparisons; including Full-Service versus Usual Services (labeled throughout as T1-C), Basic-Service versus Usual Services (labeled throughout as T2-C), Full-Service versus Basic-Service (labeled throughout as T1-T2), and the combined treatments (Full and Basic) versus Usual Services (labeled throughout as AT-C).

The analysis focused on five outcome comparisons that policymakers believed the treatment packages could positively influence the lives of recently denied disability applicants and delay the need for disability benefits. The outcomes include impacts on reported employment and earnings, perceptions of health (both mental and physical), perceptions of quality of life, reported use of health services, and disability awards. Four of the five outcomes have multiple dimensions that offer a more precise assessment of potential treatment benefits.

The initial outcome of interest was the employment rate, defined as the percentage of participants who affirm during any **one** of up to 12 follow-up interviews that they were working at a job for pay. The employment rate has been the principal outcome measure in prior assessment of the IPS model of SE services and, in this study, was the main element of the Full-Service and Basic-Service treatment packages.

The employment rate does not by itself provide sufficient information for cost-benefit purposes or to assess the value of the IPS service model against important policy-relevant standards, such as earning a living wage, earning above SGA, and working enough hours to meet basic needs. To assess the study results in a rigorous policy analysis and against these policy-relevant standards, information about additional dimensions of employment impacts is required. To provide a more complete and policy-relevant picture of the SED treatment impacts on the lives of recently denied disability applicants, we estimated impacts for five different types of outcomes. Thus, it is important to recognize the value of assessing additional dimensions to the impact analysis, such as earnings, physical health, mental health, benefit awards, and persistence with appeals. Table 4-1 presents the dimensions of interest for all five outcome types, including specific definitions for analytic purposes, and the sources of the data used for measuring these outcomes.

Table 4-1. Outcome measures, definitions, and data sources

Impact measure	Definition	Data source
Employment		
Employment rate	Percentage of participants who affirm they were working during the previous quarter on any single quarterly Interview	Participant interviews
Weeks employed	Number of weeks of participant-reported employment	Participant interviews
Total hours worked	Sum of reported hours worked across all jobs	Participant interviews
Total earnings	Sum of earnings based on reported work and pay	Participant interviews
Earnings in the past month	Sum of earnings reported in the most recent month	Participant interviews
SSA Disability Benefits		
Time to award	Number of days to disability award	SSA administrative data
Benefit appeal attempts	Number of appeal attempts during study	SSA administrative data
Total disability payments	Total disability award payments during study period	SSA administrative data
Allowance rate	Percentage of participants accepted onto the disability rolls during the study period	SSA administrative data

Table 4-1. Outcome measures, definitions, and data sources (continued)

Impact measure	Definition	Data source
Health Status		
Mental Health (SF-12)	Mental Component Score (MCS) difference scores (Study Entry vs. Exit Interview)	Participant interviews
Physical Health (SF-12)	Physical Component Score (PCS) difference scores (Study Entry vs. Exit Interview)	Participant interviews
Colorado Symptom Index	Difference scores (Study Entry vs. Exit Interview)	Participant interviews
Quality of Life		
Satisfaction with life	Difference scores (Study Entry vs. Exit Interview)	Participant interviews
Utilization of Services		
ER visits	Total emergency room visits during study	Participant interviews
Hospital overnight stays	Total overnight hospital stays during study	Participant interviews
Number of nights spent in hospital	Total number of nights spent in hospital as an inpatient during study enrollment	Participant interviews
Outpatient visits	Total outpatient visits during study	Participant interviews
Routine mental health visits	Total routine mental health visits reported in last month of study enrollment	Participant interviews
Routine general health visits	Total general health visits in last month of study enrollment	Participant interviews
Routine employment support visits	Total employment support visits in last month of study enrollment	Participant interviews

The organization of the impact results is similar across all five outcomes. First, we present overall impacts for each outcome for the entire study period and by year. Second, we assess potential predictors of the outcome. Finally, we assess the overall performance of demographic subgroups of policy interest across the three study arms.

4.2 Employment and Earnings

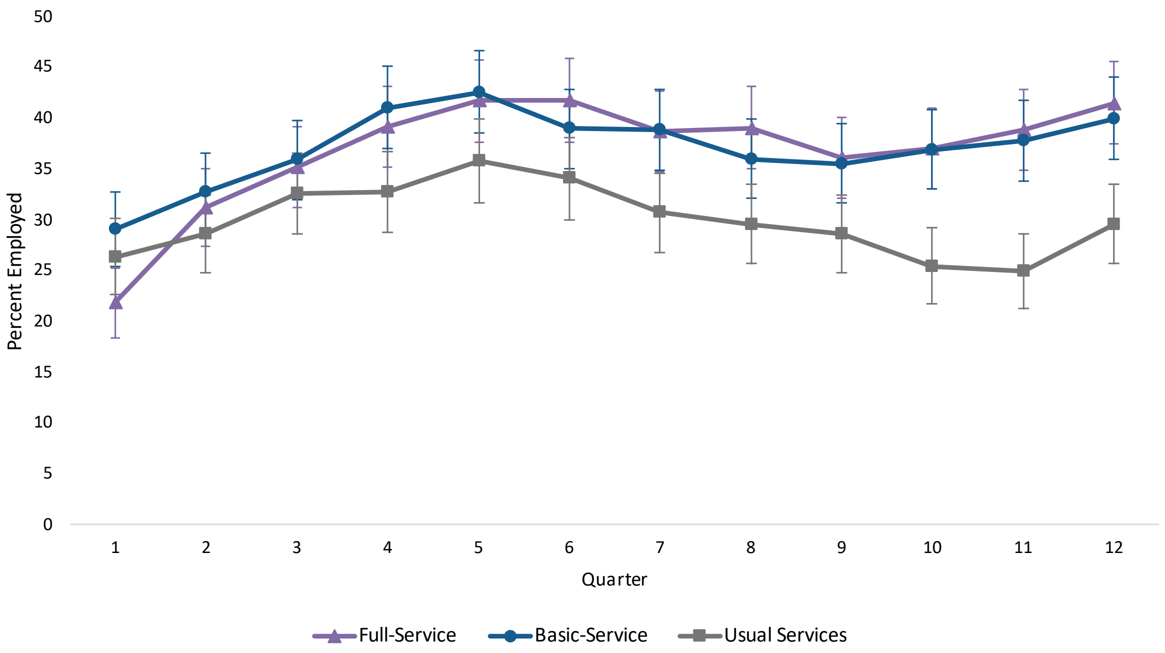
The employment and earnings focus of the impact analysis begins with a presentation of the employment rate over all 3 years combined and for each of the 3 years of study participation, followed by a presentation of the results for each additional employment dimension of interest, including *earnings*, *weeks employed*, *total hours worked*, *earnings in the past month*, and *percent earning SGA in the past month*. We then present results for only those individuals who reported working at least one job during the study, eliminating those who did not attain any employment. We also analyze the performance of study participants who worked at least half of the year over their 3 years of participation as well as those who worked at least half the year within each year of study participation.

In the second section, we present analyses of predictors of employment and earnings by demographic subgroups, including the policy-relevant subgroups associated with age, gender, race/ethnicity, and education.

4.2.1 Employment Rate

Employment Rate. We define employment rate as the percent of interview respondents who affirm they were working at any time during enrollment. Exhibit 4-2 shows progression of the employment rate with 95 percent confidence bars by study arm over the 12 quarters of study participation.

Exhibit 4-2. Employment rates by study arm across 12 quarters



Notes: Error bars represent 95 percent confidence intervals. Percentages are weighted. Includes participants who completed the final survey and did not have a gap in surveys of more than 1 year during the study period (Full-Service n = 582; Basic-Service n = 599; Usual Services n = 541).

The quarterly employment rates range from a low of about 22 percent for the Full-Services treatment group in quarter 1 to a high of about 42 percent in quarters 5, 6, and 12 in either the Full-Services or Basic-Services treatment groups. The Usual Services control group reached its highest quarterly employment rate in quarter 5 with about 36 percent. Apart from quarter 1, the Usual Services control group had lower employment rates than either the Full-Services or Basic-Services treatment groups throughout the remaining 11 quarters of the study. After the initial upward trend in employment rates over the first five quarters of the study, the two treatment groups began a slight decline in employment rates until the beginning of Year 3 (quarter 9). In Year 3, the quarterly employment rates began to rise slightly but consistently until the end of the study. The Usual Services control group followed the same pattern but with lower employment rates than either treatment group.

Table 4-2 presents the impact analysis of the overall and annual employment rates by study group. The overall employment rate reflects employment in any year of participation, while the annual rate is the percentage of participants working during each year of study participation. The overall and annual weighted percentages of study participants with employment and regression-adjusted estimates of impact between the study groups appear in Table 4-2.

Table 4-2. Employment rates during SED enrollment

	Weighted percentages						Regression-adjusted estimates of impact			
	Full-Service (T1) N = 582		Basic-Service (T2) N = 599		Usual Services (C) N = 541					
	n	%	n	%	n	%	T1-C	T2-C	T1-T2	AT-C
Employment rate in 3-year study period	427	**74.0**	443	**74.4**	345	**64.1**	8.95***	7.85***	1.11	8.41***
Employment rate in Year 1	311	**53.9**	347	**58.1**	271	**50.1**	3.49	5.72**	-2.22	4.60**
Employment rate in Year 2	345	**60.0**	344	**58.1**	264	**48.6**	10.53***	6.73**	3.8	8.64***
Employment rate in Year 3	312	**54.3**	315	**53.2**	231	**42.9**	11.33***	8.83***	2.5	10.08***

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

The overall employment rates are larger than those of any single year. The rates range from a high of 74.4 percent among participants in the Basic-Services treatment group (T2) to a low of 64.1 percent for participants in the Usual Services control group (C). The Full-Services treatment group (T1) was slightly below the T2 group with an overall employment rate of 74.0 percent. The regression-adjusted estimates of impact show positive and significant impacts of both treatment groups over the control group, with significant differences in three of the four comparisons (p < .01). While the treatment groups (T1 and T2) significantly outperformed the control group (C) both individually (T1-C; T2-C) and when combined (AT-C), the small difference in the overall employment rate between the two treatments (T1 and T2) was not significant; suggesting no meaningful difference in employment rates between the more expensive Full-Services treatment package over the less expensive Basic-Services package.

The annual employment rates reveal different patterns for the study groups. The annual employment rate for the Full-Services (T1) group was lowest in year 1 (53.9%) but increased in year 2 (60%) and remained above Year 1 in Year 3 (54.3%). The Basic-Services (T2) and Usual Services (C) groups were similar in their pattern of performance across study years with both groups achieving their highest employment rates in Year 1 (58.1% and 50.1%, respectively) and their lowest employment rates in year 3 (53.2% and 42.9%, respectively). In Year 2, the Basic-Services treatment group maintained the same annual employment rate as in Year 1 with 58.1 percent employed. However, the employment rates for the control group (C) steadily declined from the high in Year 1 of 50.1 percent to 48.6 percent in year 2 and 42.9 percent in year 3.

The regression-adjusted estimates shown in Table 4-2 reveal that the annual employment rates among the three study groups behaved quite similarly to those of the overall employment rates. In Year 1, the employment rate for the Full-Services (T1) group was not significantly different from that of the control (C) group. However, the T2-C and AT-C employment rate comparisons were both positive and significant, suggesting that the Basic-Services group significantly outperformed the control group (C). In Years 2 and 3, the differences in employment rates between treatments (T1 and T2) and the control group (C) were positive and significant (p < .01 or p < .05) both individually and when combined. As in the overall employment rates, observed differences between the employment rates of the two treatment groups (T1-T2) were not significant, again suggesting that the more expensive Full-Services treatment package did not produce better employment rates than did the Basic-Services treatment package.

Also note that the positive differential in employment rates between the treatment and control groups clearly increased over the study period. For the weighted percentages, the differential from Year 1 to Year 2 was large and then held steady between Years 2 and 3; for the regression-adjusted impacts, the differential increased over the whole study period.

4.2.2 Key Dimensions of the Employment Outcome

Additional employment outcome dimensions provide a more complete picture of the value of the employment in terms of policy-relevant standards. Chief among these dimensions are *total earnings*, *weeks employed, total hours worked*, *earnings in the past month*, and *percent earning SGA in the past month*. Table 4-3 provides the weighted measures for study participants (working and nonworking) for each study arm and the four key comparisons among the study arms overall and by study year. The regression-adjusted estimates of impact appear in the right-hand columns of Table 4-3; the full regression results (marginal effects for the covariates in the models) appear in Appendix C.

Total earnings is the sum of reported earnings weighted for the individual and averaged across all study participants. We report total earnings over the 3-year study period and annually for each of the 3 years participants were in the study. The average weighted total earnings overall for the 3-year study period ranged from a low of $13,547 for group C to a high of $17,925 for group T1. The average weighted total earnings for group T2 were $17,556. These average total earnings represent positive and significant differences between both treatment groups and the Usual Services (control) group (T1-C = $4,265, p < .01; T2-C = $3,027; p < .05) as well as between the combined treatment groups and the Usual Services (control) group (AT-C = $3,634; p < .01). Among the three annual total earnings comparisons, no significant impacts appeared in Year 1, and only a single treatment effect between T1 and C appeared in Year 2 ($1,146, p < .05). The strongest and clearest treatment effects occurred late in participation in Year 3. Table 4-3 shows both positive and significant Year 3 differences between groups T1 and T2 and group C ($2,725; p < .01; $1,978, p < .01, respectively) as well as between groups AT and C ($2,344; p < .01). As with the employment rate, there were no significant differences in weighted total earnings between the two treatment groups, T1 and T2.

Table 4-3. Key dimensions of the employment outcome by study group among all participants

	Weighted measures			Regression-adjusted estimates of impact			
	Full-Service (T1) N = 582	Basic-Service (T2) N = 599	Usual Services (C) N = 541	T1-C	T2-C	T1-T2	AT-C
Total earnings in 3-year study period	$17,925	$17,556	$13,547	$4,265***	$3,027**	$1,237	$3,634***
Total earnings Year 1	$4,145	$4,415	$3,661	$394	$426	-$33	$410
Total earnings Year 2	$6,290	$6,116	$5,083	$1,146**	$604	$542	$868*
Total earnings Year 3	$7,491	$7,025	$4,803	$2,725***	$1,978***	$747	$2,344***
Weeks employed in 3-year study period	45.92	46.53	36.98	8.98***	7.40***	1.59	8.18***
Weeks employed Year 1	12.25	13.68	11.60	0.46	1.21	-0.75	0.84
Weeks employed Year 2	16.78	16.53	13.85	2.88***	1.82*	1.06	2.34**
Weeks employed Year 3	16.89	16.32	11.53	5.63***	4.32***	1.31	4.96***
Total hours worked in 3-year study period	1,289	1,308	1,045	245.9***	184.0**	61.8	214.1***
Total hours worked Year 1	331	352	306	19.3	16.1	3.2	17.6
Total hours worked Year 2	453	465	394	56.6	35.6	21.0	45.8
Total hours worked Year 3	504	490	345	168.8***	130.5***	38.3	149.2***
Earnings in the past month (quarter 4)	$389	$359	$317	$67	$41	$25	$54
Earnings in the past month (quarter 8)	$449	$438	$337	$137***	$74	$62	$103**
Earnings in the past month (quarter 12)	$553	$590	$395	$162***	$177***	-$15	$169***
Percent earning SGA in past month (quarter 4)	4.02%	3.49%	2.85%	1.09	0.10	0.99	0.60
Percent earning SGA in past month (quarter 8)	4.40%	5.75%	4.04%	0.42	1.61	-1.19	1.02
Percent earning SGA in past month (quarter 12)	8.96%	8.42%	5.94%	2.93*	1.84	1.09	2.38*

Notes: *** p <. 01, ** p <. 05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal unweighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT=All Treatment (T1 and T2 participants combined).

Weeks employed refers to the number of weeks a participant reported having a job throughout the study. The weighted average number of weeks employed range from a low of 36.98 weeks for group C to a high of 46.53 weeks for group T2. The impact analysis is nearly identical to both the *employment rate* and *total earnings* with both treatment groups (AT) significantly outperforming the control group (C) and no significant differences between the two treatment groups (T1 and T2).

Similarly, there were few notable differences in the weighted number of weeks worked in Year 1. In Year 2 the positive differential for treatment over controls became larger and for the regression statistically significant, and it increased further in Year 3 with both T1 and T2 significantly outperforming the control group. It is worth noting that in Year 1 all three outcomes (employment rate, total earnings, and weeks employed) showed a positive difference of the Basic-Services treatment over the Full-Services treatment, though none of those differences was significant.

Total hours worked reflects the number of hours reported across all jobs during the 3-year study period. Some participants may have held more than one job at the same time or, more likely, some may have held different jobs over the study period as they attempted to regain steady employment. The number of hours worked ranges from 1,045 in group C to a high of 1,308 in group T2. Group T1 averaged 1,289 hours worked. The pattern of hours worked increases steadily over the 3 study years for both groups T1 and T2, with the Basic-Services group (T2) holding a slight edge over the Full-Services group (T1) in Years 1 and 2. However, in Year 3 group T2 had the higher average hours worked. The impact of average weighted total hours worked follows the same time pattern as the above three outcomes, with the treatment groups having a positive advantage over the control group overall and increasing from year to year. The main deviation from the above pattern where there were impacts in Year 2, total hours worked resulted in no significant differences among any of the four comparisons at either Year 1 or Year 2. While all differences in these years were positive, none of them was statistically significant.

Earnings in the past month refers to reported earnings in the past month on each of the three annual interviews (quarter 4, 8, and 12 interviews). There is no possibility of combining the three reports into an overall study measure. The average weighted past month earnings increase in each year for participants in all three study arms (T1, T2, and C). The data reveal no significant difference among any of the four comparisons in year 1 (quarter 4) but do show positive and significant differences in Years 2 and 3. In Year 2 (quarter 8) the T1 group shows a positive and significant difference in past month earnings over C, the control group ($137, $p < .01$). As well, the average past month earnings for the combined treatment groups (AT) are positive and significant compared to the control group ($103, $p < .05$). In Year 3 (quarter 12), we see a familiar pattern of positive and significant differences between groups T1 and T2 and group C ($162, $p < .01$ and $177, $p < .01$ respectively), and between AT and C ($169, $p < .01$).

Percent earning SGA in the past month refers to the percent of all study participants who reported monthly earnings that matched or exceeded SGA. Over the 3 years, the percent earning at or above SGA ranged from a low of 2.85 percent in group C at the end of Year 1 (quarter 4) to a high of 8.96 percent in group T1 at the end of Year 3 (quarter 12). While the percent of participants earning at or above SGA increased with each year for all three study groups, only two contrasts (T1-C and AT-C) approached significance (T1-C 2.93, $p < .10$ and AT-C 2.38, $p < .10$, respectively).

Participants Who Worked at Least One Job. Table 4-3 shows the key dimensions of employment in the context of the entire analytic sample, including large numbers of participants who did not get a job during any year of the 3-year study period. While these total sample statistics are critical in understanding the impact of the treatment conditions on the population, the large numbers of participants with no job obscures somewhat the average performance of workers across most of the key dimensions because the averages include so many zeros. The larger the proportion of participants who do not get jobs certainly reduces the outcome values (e.g., employment rate, earnings). However, for some outcomes, such as earnings in the past month, the inclusion of participants who did not get a job at all or during a particular year of study participation distorts interpretations of results for those who did hold a job when trying to understand information about the jobs taken by participants. For this reason, Table 4-4 presents the performance data of only those participants in each study arm who held at least one job in each year of the study.

A review of the data in Tables 4-3 and 4-4 show that removal of nonworkers from the data substantially increases the magnitudes of all the weighted measures, especially the annual data. For example, the Year 3 total earnings of the Usual Services control group doubled in each study year (from $3,661 in Year 1 to $7,306; from $5,083 in Year 2 to $10,454, and from $4,803 in year 3 to $11,210) with the removal of nonworkers. These new weighted measures provide an opportunity to view the jobs taken by study participants in the context of external standards such as part- and full-time work, poverty levels, and SGA.

Annual total earnings among the Full-Services (T1) workers averaged $7,687 and was similar for Basic-Services (T2) and control group (C) workers. The weighted average total earnings among group T1 workers increased 36 percent to $10,477 in Year 2, 39 percent to $10,531 among group T2 workers, and 43 percent to $10,454 among group C workers. In Year 3, the highest average total earnings among all workers, the average total earnings among T1 workers increased another 32 percent to $13,798, and among T2 workers another 25 percent to $13,203, but only 7 percent to $11,210 among group C workers.

A review of the regression-adjusted estimates of impact Table 4-3 shows that nearly all of the significant differences in average total earnings noted go away with removal of nonworkers from the analysis. While differences between the treatment groups and the control group, both individually and combined, were previously positively and statistically different (see Table 4-3), many of them are not with removal of nonworkers from the analysis (see Table 4-4). The overall earnings impact is only marginally significant for the combined treatment groups but becomes significant only in Year 3, when the regression-adjusted Full-Services impact (T1-C) is $2,588 ($p < .05$), and the combined treatment group impact (AT-C) is $2,290 ($p < .05$). In Year 3, Basic-Services impact (T2-C) approaches significance with a value of $1,998 ($p < .10$). The decline in significance of the earnings impacts is statistically consistent with the decline in sample size when the large number of nonworkers is excluded. The year-to-year pattern of impacts is, however, altered with Year 2 showing little or no increase compared to Year 1. This suggests the possibility that much of the Year 2 earnings impacts for the entire study population was due to employment gains while relatively more of the very large and significant Year 3 earnings impact reflected the ability of participants who gained employment in Years 1 or 2 to maintain those job and perhaps even secure pay increases as they gained on-the-job experience.

Of course, the smaller overall earnings impacts noted here, especially for Years 1 and 2, does not diminish the importance of the impacts seen in Table 4-3. Those impacts reflect the total study population. The impacts noted here are limited to the **working** subgroup of the study population.

The *weeks employed, total hours worked, and earnings in the past month* data reveal very similar patterns to changes in the *total earnings* data when the analysis excludes nonworkers. Given that the Year 3 data is consistently highest for all dimensions of employment, it is instructive to note that both total earnings and monthly earnings in Year 3 begin to approach poverty-level statistics. The poverty earnings indicator for a single person in 2022 was a monthly income of $1,133 or an annual income of $13,500. Note that the earnings in the past month for quarter 4 (Year 3) ranges from $922 in the control group to $1,113 in the Basic-Services group. Further, average total earnings ranges from $11,210 in the control group to $13,798 in the Full-Services treatment group. The data in Table 4-2 shows the study treatment packages were equally significantly better than Usual Services in yielding higher employment rates. Table 4-3 shows those higher employment rates resulted in higher earnings, weeks worked, hours worked, and recent monthly earnings across all study participants. The earnings trajectory of participants in the two treatment groups appears to be quite positive. Table 4-4 shows that by Year 3, earnings and recent monthly income begins to approach poverty levels.

Finally, Table 4-4 show that the percentage of workers earning SGA in the past month ranges from a low of 5.69 percent in the Usual Services group in Year 1 to a high of 16.51 percent in the Full-Services treatment group in Year 3. The percentage of workers earning SGA in the past month increases each year in all three study arms. Further, the data reveal no significant differences between any of the study arms in any year among the working subpopulation.

Table 4-4. Key dimensions of the employment outcome by study group among those who worked at least one job during the study period

	Weighted measures			Regression-adjusted estimates of impact			
	Full-Service (T1)	Basic-Service (T2)	Usual Services (C)	T1-C	T2-C	T1-T2	AT-C
Total earnings in 3-year study period	$24,216	$23,588	$21,123	$3,093	$2,465	$628	$2,775*
Total earnings Year 1	$7,687	$7,598	$7,306	$380	$292	$88	$334
Total earnings Year 2	$10,477	$10,531	$10,454	$77	$23	-$54	$50
Total earnings Year 3	$13,798	$13,203	$11,210	$2,588**	$1,993*	$596	$2,290**
Weeks employed in 3-year study period	62.04	62.52	57.66	4.38	4.86	-0.08	4.62
Weeks employed Year 1	22.73	23.54	23.16	-0.43	0.38	-0.81	-0.01
Weeks employed Year 2	27.95	28.46	28.48	-0.53	-0.01	-0.52	-0.27
Weeks employed Year 3	31.11	30.68	26.91	4.19***	3.76**	0.43	3.98***
Total hours worked in 3-year study period	1,741	1,757	1,630	112	127	-15	119
Total hours worked Year 1	615	606	611	4	-4	8	-0.4
Total hours worked Year 2	755	801	811	-56	-10	-45	-33
Total hours worked Year 3	929	922	805	124	116	8	120**
Earnings in the past month (quarter 4)	$702	$614	$613	$89	$1	$88	$42
Earnings in the past month (quarter 8)	$744	$749	$697	$47	$52	-$5	$49
Earnings in the past month (quarter 12)	$1,029	$1,113	$922	$107	$191*	-$84	$149*
Percent earning SGA in past month (quarter 4)	7.45%	6.00%	5.69%	1.76	0.31	1.45	1.00
Percent earning SGA in past month (quarter 8)	7.33%	9.89%	8.31%	-1.00	1.58	-0.26	0.28
Percent earning SGA in past month (quarter 12)	16.51%	15.53%	13.87%	2.60	1.70	1.00	2.15

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal unweighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT=All Treatment (T1 and T2 participants combined).

Steady Workers. In the context of the study, we defined steady work as working at least half the time during the 3-year study period or at least half the time during a single year. Table 4-5 presents a summary of the percentages of workers in each study arm who are designated as steady workers over the study period (working at least 78 weeks of the 156 weeks) and in each year of the study (working at least 26 of the 52 weeks of each year). The percentages in Table 4-5 reflect only the subgroup of participants in each study arm who worked at least one job during the 3-year study period. The table shows that over the entire 3-year study period more than a third of workers were designated *steady workers,* with both treatment groups yielding higher percentages than the control group (T1, 37.26; T2, 37.17; C, 30.97). The differences were both positive and significant between each treatment group and the control group (T1-C = 7.85, p < .05; T2-C = 6.48, p < .05) and when combined (AT-C = 7.17, p < .05).

Table 4-5. Percentage of study participants who worked at least half of the time by study arm and year of enrollment, among those who worked at least one job

	Weighted percentages			Regression-adjusted estimates of impact			
	Full-Service (T1)	Basic-Service (T2)	Usual Services (C)				
	%	%	%	T1-C	T2-C	T1-T2	AT-C
Study period	37.26	37.17	30.97	7.85**	6.48**	1.38	7.17**
Year 1	38.39	38.24	38.15	0.25	<0.01	0.15	0.17
Year 2	51.12	52.34	49.32	1.79	3.02	-1.22	2.40
Year 3	58.66	56.06	47.53	11.13**	8.53*	2.59	9.82**

Notes: ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal unweighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT=All Treatment (T1 and T2 participants combined).

Annual weighted percentages across the three study groups provide more clarity in understanding the steady worker population in the study. In Years 1 and 2, the steady worker percentages were remarkably similar across the three study groups. In Year 1 the percentages in all three groups were just over 38 percent; T1 was 38.39, T2 was 38.24, and C was 38.15. Year 2 shows similar increases across all three groups (51.12, 52.34, and 49.32 respectively). These data resulted in no significant differences among the three study arms. However, between Year 2 and Year 3, the percentages of steady workers increase 7.54 percentage points for T1 to 58.66 percent and 3.72 percentage points to 56.06 percent for T2. The control group (C) decreased 1.79 percentage points from Year 2 to Year 3 from 49.32 percent to 47.53 percent. Differences in these weighted percentages were positive and significant between T1 and C (11.13, p < .01) and between the combined treatments and the control group (AT-C= 9.82, p < .01). The T2-C comparison approached significance (T2-C=8.53, p < .10). At no time were the positive or negative differences between the two treatment arms significantly different.

4.2.3 Predictors of Employment

The SED study population comprised of recently denied disability applicants applied for disability because they believed they could no longer work or work at a level that provided a living for them. As noted above, many individuals in all three study arms did return to work at some point in the study, with many of them entering and keeping a long-term job (over half worked at least half-time by Year 3), as indicated in Table 4-5. We wanted to better understand what factors or characteristics associated with study participants increased or decreased the chance of working again. We completed a logistic regression with the dependent dichotomous variable *worked/did not work* at any time during the 3 years of study participation. Working from a relevant mix of demographic, education, health, disability, and community context variables, we attempted to determine which, if any, variables might predict higher (or lower) chances of getting a job.

Table 4-6 contains the results of the logistic regression (pseudo $R^2 = .1894$) including the variable labels, marginal effects, standard errors, and p-values for each variable. The data reveal nearly a dozen variables that significantly affected chances of getting (or not getting) a job among the population participating in the SED. First, and as might be expected from results presented earlier, having been randomized to the *Full-Services* or *Basic-Services* treatment groups (8.954, $p < .000$; 7.848, $p < .000$, respectively) predict having a job compared to the referent of assignment to the Usual Services study arm. This result strongly suggests that persons like those in the study would be more likely to get a job if they received services like those offered in the Full-Service or Basic-Service treatments. From a health and disability perspective, having better mental health (reported at baseline) as measured by the SF-12 MCS (0.308, $p = .015$) and physical health (reported at baseline) as measured by the SF-12 PCS (.344, $p= .008$) were significant in predicting work, as were indicators from the Work Disability Functional Assessment Battery (WD-FAB) of higher *upper body functioning* (.931, $p < .003$) and greater *community mobility* (3.153, $p = .017$).

Table 4-6 also reveals a number of indicators that predict a less likely chance of getting a job among SED study participants. Chief among these variables were conditions people reported at baseline including *not working at baseline but worked in the past 2 years* (-21.371, $p < .001$), *worked but not in the past 2 years* (-35.219, $p < .001$), and *never worked* (-38.537, $p < .001$). All three of these conditions reported by SED enrollees showed their chances of getting work were between 21 and 39 percent less likely to get a job at any time during the study period compared to those participants who were working at baseline. The model also identified two additional negative predictors of work during the study period. They were *age* (-.643, $p < .000$) and *TANF receipt in the past 12 months* (-8.731, $p < .015$). For each additional year of age, the data suggest there is a .64 percentage point decrease in the likelihood of any employment. Also, persons who received TANF in the past 12 months (reported at baseline) were less likely to get a job during the study period.

Table 4-6. Logistic regression results of predicting any employment during 36-month study period

Variable	Marginal effect[a]	Standard error	p-value
Study arm			
Reference group: Usual Services	–	–	–
Basic-Service dummy	7.848	2.508	0.002
Full-Service dummy	8.954	2.450	0.000
Work status at baseline			
Reference group: Currently working	–	–	–
Not currently working but worked in prior 2 years	-21.371	2.128	0.000
Not currently working but worked prior to the past 2 years	-35.219	3.503	0.000
Never worked	-38.537	7.932	0.000
Missing	-7.092	13.878	0.609
Weeks employed (past 2 years baseline)	0.087	0.064	0.176
Total earnings (past 2 years baseline)	< 0.001	< 0.001	0.366
Age	-0.643	0.124	0.000
Female	-0.125	2.151	0.954
Race and ethnicity			
Reference group: White non-Hispanic	–	–	–
Black non-Hispanic	1.687	2.549	0.508
Hispanic	3.891	3.197	0.224
Two or more races non-Hispanic	-3.425	3.813	0.369
Other or missing	4.456	5.856	0.447
Education			
Reference group: Less than high school	–	–	–
Completed high school	-2.275	2.989	0.447
Some college or technical school	-1.180	2.958	0.690
Associate degree	1.794	4.523	0.692
Bachelor’s degree or better	-1.657	4.341	0.703
Nights spent in jail prior to enrollment	0.077	0.052	0.140
SNAP receipt in 12 months prior to enrollment	-0.867	2.343	0.712
TANF receipt in 12 months prior to enrollment	-8.731	3.604	0.015
MCS at baseline	0.308	0.127	0.015
PCS at baseline	0.344	0.130	0.008
Urban	4.428	3.200	0.166
WD-FAB Domains			
Basic Mobility	-0.563	0.328	0.087
Upper Body Function	0.931	0.313	0.003
Communication and Cognition	-0.208	0.199	0.296
Resilience	0.251	0.135	0.062
Mood and Emotions	-0.235	0.128	0.067
Interpersonal Interactions	0.149	0.144	0.302
Community Mobility Drive	3.153	1.325	0.017
County Average Weekly Wage – Private (BLS)	0.002	0.004	0.529
State Average Hospitalization Rate Due to COVID-19 (CDC)	-0.179	0.357	0.617
County Average Unemployment Rate (LAUS)	-0.879	0.956	0.358

Note: [a] Marginal effects represent average percentage point changes in the likelihood of employment during the study period. Pseudo R-squared = 0.1894.

4.2.4 Employment Rates by Policy-Relevant Subgroups

Demographics are increasingly relevant to health and human services across the United States. Whether it involves access to services or the actual provision of services themselves, being a member of a minority group, older, female, or poorly educated affects the types and intensity of services one receives. While this report has clearly established the impacts of the Full-Services and Basic-Services treatments on employment rates and earnings, we also wanted to assess how particular demographic subgroups fared in order to explore differences among the study arms. Therefore, we analyzed employment rates by age, gender, race/ethnicity, and education to determine whether the patterns of employment success among subgroups were the same or different.

Table 4-7 presents weighted percentages of overall employment by subgroup and study arm. As noted earlier, we define employment in this context as having a job at any point during the 3-year study period. The weighted percentages represent the adjusted percentage of participants who held a job representing that subgroup (the number of participants comprising the subgroup is not shown). The weighted percentage of employed persons in the 18-34 age group within the Full-Service treatment was 82.0 percent. Not shown in the table is the actual percentage of participants who represent the total subgroup membership (i.e., including nonworkers). Thus, the weighted percentages for each subgroup within the variable (e.g., gender) are independent of one another. For example, the weighted employment percentages in the Full-Services study arm for males is 78.5 percent and 71.2 percent for females.

Ideally, we expect patterns among the various subgroups that look similar to the pattern shown in the top row of Table 4-7, showing the overall employment rate within each study arm and the four relevant regression-adjusted estimates of impact. Naturally, smaller sample sizes will alter the specific percentages and estimates, but our expectation is the general patterns of impact in each subgroup (i.e., each row of the table) should be similar to the overall pattern.

Employment Rates by Age. We know from the analysis above that *age* is a predictor of employment among this study population. As noted above, for each additional year of age, the percentage of participants employed drops by around 0.64 percent. This statistic is consistent with the data represented by the two age subgroups: participants in the age 18-34 subgroup, and participants in the age 35-49 subgroup. Across all three study groups, the participants in the younger age group substantially outperformed participants in the older age group across the three study groups (82.0% vs. 68.0%, 80.9% vs. 69.3%, and 70.7% vs. 59.5%). The patterns of impact for the two age groups are similar but not exactly alike. Both age groups in the Full-Services treatment significantly outperformed their respective counterparts in the Usual Services (control) group (T1-C, 8.40, $p < .05$; and T1-C, 8.90, $p > .01$). We see the same result for the two age groups with the combined Full-Service and Basic-Service treatment groups versus the Usual Services (control) group (AT-C). Also consistent across the two age groups was the fact that the two treatment groups (T1-T2) yielded no differences. The Basic-Services treatment by itself was not significantly different from the Usual Service group (T2-C) for the 18-34 age group, but it was statistically significant for the 35+ age group (T2-C, 9.29, $p < .01$), suggesting that older study participants gained more from Basic-Service than did younger participants.

Table 4-7. **Overall employment rates by policy-relevant subgroups**

	Weighted percentages			Regression-adjusted estimates of impact			
	Full-Service (T1) N = 582	Basic-Service (T2) N = 599	Usual Services (C) N = 541				
	%	%	%	T1-C	T2-C	T1-T2	AT-C
Employment rate in 3 year study period	74.0	74.4	64.1	8.95***	7.85***	1.11	8.41***
Age							
Age 18-34	82.0	80.9	70.7	8.40**	5.38	3.03	6.94**
Age 35+	68.0	69.3	59.5	8.90***	9.29***	-0.39	9.07***
Gender							
Male	78.5	77.1	64.3	12.62***	10.94***	1.68	11.75***
Female	71.2	72.5	64.0	6.39*	5.53	0.85	5.98**
Race and ethnicity							
White non-Hispanic	72.6	69.9	66.4	3.97	1.72	2.25	2.94
Black non-Hispanic	78.8	76.3	64.5	13.69***	8.93**	4.76	11.11***
Hispanic	79.9	77.3	63.1	17.05**	15.37**	1.68	16.09**
Two or more races non-Hispanic	65.7	81.8	52.8	17.64**	28.80***	-11.16	22.26***
Other/missing	64.1	83.8	63.2	-11.91	2.16	-14.07	-5.65
Education							
Less than high school	70.1	70.6	62.2	11.61*	10.43	1.18	11.07*
Completed high school	75.4	75.7	58.7	11.85***	11.73***	0.11	11.78***
Some college or technical	74.3	74.5	64.2	9.98**	6.75	3.23	8.47**
Associate degree	74.4	71.6	80.0	-11.95	-8.49	-3.46	-10.09
Bachelor's or better	76.5	78.1	73.7	4.92	5.46	-0.54	5.20
Working at enrollment							
Yes	98.1	96.5	94.1	3.96	1.74	2.22	2.86
No	68.3	68.0	57.7	10.70***	9.58***	1.12	10.14***
On probation in year prior to enrollment							
Yes	79.0	74.3	75.0	3.56	-5.16	8.72	-1.03
No	73.4	74.3	62.8	9.63***	9.46***	0.17	9.52***

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Unweighted counts appear in Appendix C. Weighted percentages do not necessarily equal unweighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT=All Treatment (T1 and T2 participants combined).

Employment Rates by Gender. The weighted percentage pattern for males is the same as that for the overall group shown in the top row of the table. For females, only the combined treatment groups when contrasted with the control group reveals a significant difference (AT-C = 5.98, p < .05). While the T1-C contrast for females is positive, the difference between the two study arms is smaller and only approaches significance (6.39, p < .10). Where the T2-C contrast was statistically significant for males, it was not for females.

Employment Rates by Race/Ethnicity. Table 4-7 shows that *White non-Hispanic* study participants did not achieve the same impacts of either treatment package that the overall study population achieved. As noted in the table, the only contrast for the *White, non-Hispanic* subgroup that was consistent with the overall population (row 1 of the table), was the T1-T2 contrast that also resulted in a finding of no difference. Otherwise, the data suggest that the White, non-Hispanic subgroup fared no better, or worse, than participants in the other two study arms. Participants in subgroups identifying as *Black non-Hispanic*; *Hispanic*; or persons of *two or more races* all performed consistently with the patterns of the overall study groups. While levels of significance varied between p < .01 and p < .05, the impact patterns were identical. One interesting deviation from the pattern was with persons of two or more races.

Employment Rates by Education. Among the five categories of education, only the subgroup comprising participants who *completed high school* as their highest level of education followed the same pattern as the overall study groups. Participants in both treatment groups significantly outperformed participants in the Usual Services (control) group; both treatment groups when combined also achieved a significant performance over the control group; and the table shows that, consistent with the overall population, no statistically significant difference existed between the two treatment groups. Participants in the *less than high school* subgroup showed weaker results than the *completed high school* group. Contrasts T1-C (11.61, p < .10) and AT-C (11.07, p < .10) approached significance. The T2-C contrast was not significant, and neither was the T1-T2 contrast. But no difference between the two treatments is consistent with the overall population performance.

Participants in the two subgroups comprising people with an *associate degree* and people with a *bachelor's degree or higher* showed nonsignificant performance differences across all three study groups, suggesting that participants with these degrees were not impacted by either treatment group over their counterparts or the Usual Services (control) group. Among participants with an *associate degree*, those in the Usual Services (control) group outperformed participants in both treatment groups, individually and combined, an anomaly not seen elsewhere in the subgroup analysis. The Full-Service treatment subgroup of participants with *some college or technical certificate* and the combined Full-Service and Basic-Service treatment groups significantly outperformed participants in the Usual Services (control) group (T1-C, 9.98, p < .05; AT-C, 8.47, p < .05, respectively).

Employment Rates for Those Working/Not Working at Enrollment. Nearly 20 percent of study participants were *working at enrollment*. They entered the study with the idea of getting help to keep a job or seek a better job. As one might expect, the study participants *not working at enrollment* show the exact same pattern as the overall study population. T1-C, T2-C, and AT-C are all significant beyond p < .01. Just as clear, the employment rate data associated with participants *working at enrollment* very clearly shows they did not benefit from the treatment packages in the same ways the overall study population benefited. The data show there were no significant differences for any of the four contrasts targeting impact.

Employment Rates for Those on/Not on Probation in Year Prior to Enrollment. Similar to the working at enrollment data, participants who reported being *on probation in the year prior to enrollment* also showed no benefit from the treatment packages over the Usual Services control group, as none of the four impact bearing contrasts showed a significant difference. Persons reporting that they were *not on probation in the year prior to enrollment* faired quite well in the study. The pattern of significance mirrors exactly that of the overall study population.

Total Earnings by Subgroups. Table 4-8 shows the earnings by subgroup data. As with employment rates, we expect the pattern of significance for the earnings subgroups to follow earnings impact pattern of the overall population. The top row of Table 4-8 shows the overall earnings over the 3-year study period; only four of the potential 18 subgroups follow the same pattern as the overall study population. Males, Black non-Hispanics, participants with Associate degrees, and participants who were not on probation in the year prior to study enrollment all performed as the overall study population. These subgroups showed the same three positive and significant differences between the Full-Services and Basic-Services treatment group both individually and combined over the Usual Services control group. However, all four cases of the complementary subgroups, with a couple of minor exceptions, showed no significant differences with any of the four target contrasts. The single notable exception was the subgroup identified as *two or more races, non-Hispanic*. In this case, the combined treatment groups (AC) showed a positive and significant difference in total earnings over the control group (C).

The patterns for the two age subgroups (ages 18-34; age 35+) were nearly identical but different from the overall study group. In both age subgroups, the T1-C contrast was significant (p < .05) as were the AT-C contrasts (p < .05). For participants in the age 35+ subgroup, the T2-C contrast approached significance, with a p-value < .10.

The patterns of effectiveness for the two *working at enrollment* subgroups were also very similar to each other and nearly identical with the age subgroups. Participants in the *working at enrollment* subgroup showed significant differences in total earnings between the Full-Services treatment group and the Usual Services control group (T1-C = $9,266, p < .05) and between the combined treatments and the control group (AT-C = $8,662, p < .05). Table 4-8 shows that performance of participants in the *not working at enrollment* subgroup showed the same pattern. For participants in the *working at enrollment* subgroup the T2-C contrast approached significance, with a p-value < .10.

Table 4-8. Earnings by study group by age, gender, race, and education

	Full-Service (T1) N = 582	Basic-Service (T2) N = 599	Usual Services (C) N = 541	Regression-adjusted estimates of impact			
	M	M	M	T1-C	T2-C	T1-T2	AT-C
Total earnings in 3-year study period	$17,925	$17,556	$13,547	$4,265***	$3,027**	$1,237	$3,634***
Age							
Age 18-34	$17,429	$17,933	$12,372	$4,749**	$2,918	$1,831	$3,767**
Age 35+	$18,298	$17,259	$14,383	$3,921**	$3,139*	$782	$3,536**
Gender							
Male	$22,141	$21,186	$12,653	$8,426***	$7,102***	$1,324	$7,738***
Female	$15,262	$14,908	$14,189	$1,421	$191	$1,230	$804
Race and Ethnicity							
White non-Hispanic	$18,583	$16,028	$15,247	$2,637	$90	$2,547	$1,514
Black non-Hispanic	$17,272	$18,656	$12,314	$6,195**	$5,043**	$1,152	$5,512***
Hispanic	$18,452	$18,079	$13,693	$4,256	$4,633	-$378	$4,500
Two or more races non-Hispanic	$14,820	$18,798	$9,832	$6,041	$7,240*	-$1,199	$6,590**
Other/missing	$19,798	$20,472	$8,790	$7,394	$4,943	$2,450	$6,034

Table 4-8. Earnings by study group by age, gender, race, and education (continued)

	Full-Service (T1) N = 582	Basic-Service (T2) N = 599	Usual Services (C) N = 541	Regression-adjusted estimates of impact			
	M	M	M	T1-C	T2-C	T1-T2	AT-C
Education							
Less than high school	$11,155	$10,959	$10,386	$2,225	$1,356	$870	$1,802
Completed high school	$13,110	$14,589	$10,490	$2,195	$3,193*	-$998	$2,784*
Some college or technical	$17,478	$18,599	$13,693	$3,534	$2,234	$1,300	$2,854
Associate degree	$25,705	$24,650	$13,255	$12,927***	$14,186***	-$1,259	$13,520***
Bachelor's or higher	$39,122	$28,717	$28,024	$9,957	-$76	$10,033	$5,062
Working at enrollment							
Yes	$37,002	$35,729	$28,690	$9,266**	$8,075*	$1,191	$8,662**
No	$13,436	$12,395	$10,272	$3,223**	$1,730	$1,493	$2,471**
On probation in year prior to enrollment							
Yes	$12,000	$16,656	$12,001	$2,132	$2,715	-$582	$2,484
No	$18,647	$17,555	$13,734	$4,533***	$2,902**	$1,631	$3,720***

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

4.2.5 Earnings Impacts by Enrollment Date

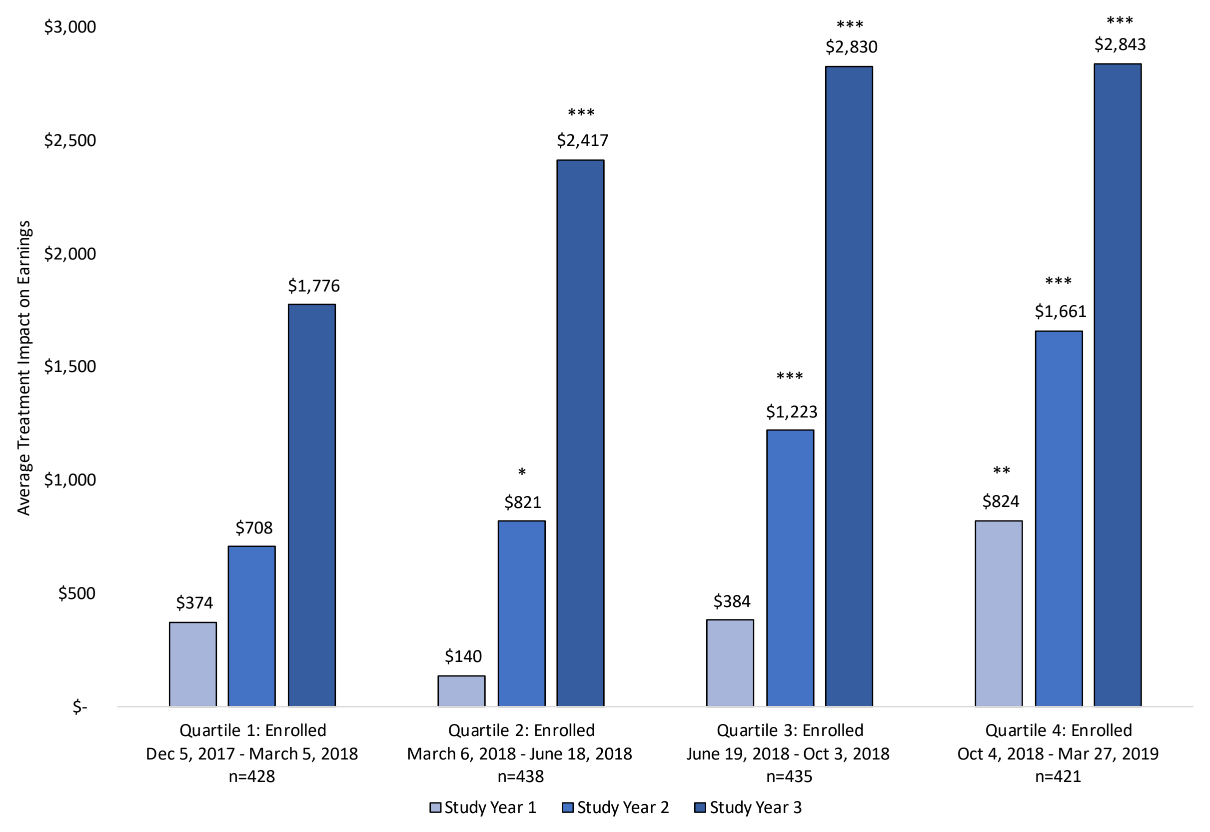
The SED recruited participants over 16 months, from December 2017 through March 2019. Due to the extended recruitment period, early enrollees may have a different experience from those who enrolled later in the process. Such differences might have occurred because of changing external circumstances (such as the emergence of the COVID-19 pandemic in early 2020), changing labor market conditions, or modifications and improvements by the site teams in their service delivery practices. To explore the possibility of different treatment effects depending on early or late enrollment, we examine differences in earnings impacts based on participants' enrollment date.

Exhibit 4-3 provides a summary of regression-adjusted impacts for all treatment group participants (AT) compared to Usual Services. The figure shows separate estimates based on enrollment dates, with participants divided into four similarly sized groups by their enrollment date: December 5, 2017-March 5, 2018 (quartile 1), March 6, 2018-June 18, 2018 (quartile 2), June 19, 2018-October 3, 2018 (quartile 3), and October 4, 2018-March 27, 2019 (quartile 4).

The results indicate that treatment participants who enrolled in quartiles 3 and 4, on average, saw larger earnings gains compared to the Usual Services control group participants who enrolled in the same quartiles compared to the early quartiles (1 and 2). The first quartile of treatment group enrollees (enrolled from December 5, 2017 through March 5, 2018) did not see significant impacts on earnings relative to the control group in their first, second, or third year of enrollment.

A possible explanation for the larger earnings impacts among later enrollees compared to early enrollees is that sites learned over time how to better serve SED participants. Early enrollees received treatment from staff who had less experience with the recruited population of denied disability applicants with mental impairments. Staff may have learned to address the specific needs of SED participants, providing more effective services to later enrollees.

Exhibit 4-3. Average earnings impacts on treatment participants by enrollment date quartiles



Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates represent the regression-adjusted difference between the combined treatment groups (AT) and the control (Usual Services) group. Full model results appear in the appendix.

Another explanation for the relative underperformance of the first quartile is that the timing of enrollment relative to the disability denial decision was longer for some early enrollees. The first group of denied applicants used for recruitment beginning in December 2017 included denials as early as July 2017. This delay resulted in a median difference of 116 days between SSA denial and enrollment in the SED for quartile 1, compared to 56, 42, and 37 days for quartiles 2, 3, and 4, respectively. The longer delay may have meant that some participants in quartile 1 were too far removed from the denial decision to benefit from the intervention.

For each of the four enrollment quartiles, earnings impacts increased from the first year of study enrollment through the third year of study enrollment. Treatment group participants in quartiles 3 and 4, for example, saw an average of over $2,800 in earnings more than the control group, compared to $1,223 and $1,661 for quartiles 3 and 4 in the second year of study enrollment, respectively.

Finally, it is important to note that changing local labor market conditions may in fact have also influenced the time trends and subgroup results discussed above. Evidence from a simple analysis allowing for interaction between the county unemployment rate over the study period and the combined treatment indicator shows a treatment effect on earnings (including the unemployment rate interaction) for the last quartile of study entrants was about 22 percent larger than for the first quartile. Further research could explore the implications of these trends for the subgroup analysis since the impacts of changing labor market conditions could have been quite different across the relevant subgroups. Since there was also a strong decrease over time in the days from claim denial to study entrance, it is also important to explore the treatment effect trends allowing for both interactions with local labor market conditions and with the days from denial to study entrance.

4.3 SSA Benefits

This section summarizes the impact of the SED on SSA benefit awards and applications. The primary measure of SSA benefit receipt is the allowance rate. We also measure the impact of the SED on secondary measures, including appeal attempts and total benefit payment amounts.

4.3.1 Allowance Rate

Among those enrolled in the SED, approximately 15 percent (439 participants) received an allowance for disability benefits (either SSDI, SSI, or both SSDI and SSI) during the 3 years of enrollment in the study. Table 4-9 summarizes the allowance rates for the entire 3-year study period and separately for each year of enrollment. For the 3-year study period, there was no significant difference in allowance rates between the study arms. Looking at each year of enrollment separately, we find no significant differences in the annual allowance rates across the three study arms.

Among the 439 participants who received allowances, approximately 36 percent (157 participants) received an allowance for SSDI only; 38 percent (167 participants) were allowed for SSI only; and a smaller percentage, 26 percent (115 participants) were allowed for both SSDI and SSI. Table 4-9 provides the percentages of participants allowed for each program separately within each study arm. Generally, there were no significant differences in the allowance rates by program type across study arms. The exception is the difference between SSDI allowance rates between the Full-Service and Basic-Service groups (Basic-Service participants were less likely than Full-Service to receive an allowance for SSDI).

Allowance rates were the highest in the second year of study enrollment. It is important to note that the timing of allowances depends on several factors, including the timing of appeals, the caseload and backlog of the local SSA field office, and the complexity of the case. Many of these factors are outside of the participant’s control. Caseloads at the local SSA field office, for example, and the specifics of the individual’s claims and medical evidence have a large influence over the amount of time needed to reach a decision.

Table 4-9. SSA benefit allowance rates

	Percentages						Regression-adjusted estimates of impact			
	Full-Service (T1) N = 976		Basic-Service (T2) N = 987		Usual Services (C) N = 981					
	n	%	n	%	n	%	T1-C	T2-C	T1-T2	AT-C
Allowance rate in 3-year study period (SSDI or SSI)	143	14.7	144	14.6	152	15.5	-0.89	-0.91	0.02	-0.90
Allowance rate in Year 1	35	3.6	38	3.9	34	3.5	0.30	0.35	-0.05	0.33
Allowance rate in Year 2	64	6.6	83	8.4	84	8.6	-2.06*	-0.02	-2.04*	-1.05
Allowance rate in Year 3	44	4.5	23	2.3	34	3.5	0.99	-1.18	2.17***	-0.10
Allowance type										
SSDI only	63	6.5	42	4.3	52	5.3	0.10	-0.98	2.00**	0.05
SSI only	48	4.9	59	6.0	60	6.1	-1.13	0.01	-1.15	-0.55
SSDI and SSI	32	3.3	43	4.4	40	4.1	-0.81	0.09	-0.90	-0.35

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

Allowances and Work. Receiving an allowance for disability benefits affects the decision to work. Many participants sought income to meet basic needs from whatever source they could (e.g., work, benefits, or a combination of work and benefits). Allowed participants may quit jobs or end job searches because they no longer needed income from work to meet their basic needs. Many participants worked while appealing or reapplying denial decisions. About 42 percent of those who were appealing or reapplying during study enrollment were employed during the study. To examine the relationship between work and allowances, Table 4-10 provides a summary of the annual employment rates separately for participants allowed in their first, second, or third year of enrollment in the SED.

The analysis provides evidence that suggests the need for further research into the relationship between disability allowance and work. There are three key takeaways. First, employment rates were lower overall for participants who were allowed in either year 1, 2, or 3 of the study compared to those who were never allowed. Second, for those who were allowed, employment rates dropped for each subsequent year. Third, although the employment rates among allowed participants was low, some allowed participants (about one in five) were employed after receiving an allowance.

Table 4-10. Employment rates among participants by allowance year

	Allowed in study year						Not allowed	
	1		2		3			
	n	%	n	%	n	%	n	%
All participants	59		153		64		1,446	
Year 1	17	30.2	53	34.2	28	44.3	831	57.5
Year 2	16	27.7	46	30.5	23	37.5	868	55.8
Year 3	11	19.5	32	21.4	15	23.9	858	50.3
Full-Service	17		43		31		491	
Year 1	3	18.6	15	35.2	15	49.3	278	56.9
Year 2	5	31.1	9	21.0	13	43.8	318	65.2
Year 3	4	25.6	8	18.4	9	28.8	291	59.8
Basic-Service	26		51		15		507	
Year 1	11	44.4	21	41.1	6	41.6	309	60.9
Year 2	9	35.4	21	41.1	5	35.5	309	61.4
Year 3	5	20.3	13	26.1	4	29.9	293	53.2
Usual Services	16		59		18		448	
Year 1	3	20.1	17	27.6	7	38.4	244	54.5
Year 2	2	11.9	16	28.3	5	28.8	241	53.3
Year 3	2	11.9	11	19.6	2	11.4	216	48.2

Notes: Percentages (%) are weighted; counts (n) are unweighted. Weighted percentages do not necessarily equal weighted counts.

4.3.2 Other Measures of SSA Benefits

Appeal Attempts

After receiving a denial letter, denied applicants have 60 days to file their first appeal. Generally, if the applicant misses the 60-day deadline for an appeal, then they must submit a new application instead of an appeal. The time until approval depends in part on the location of the participant because SSA field offices handle different numbers of applications. Cases that are more complex may require the applicant to provide additional evidence and attend a hearing with an Administrative Law Judge; these steps can add substantial time to the approval process.

The key measure of appeals is the number of appeal attempts made during the study period. To assess the impact of the SED on appeals, we measure the number of appeal attempts made during the study period for each study arm and report the differences between each study arm.

It is important to note that while the decision to appeal is up to the study participant, the total number of appeals depends in part on the timing of SSA disability decisions. That is, participants can only file an appeal when SSA issues a denial decision. If SSA needs longer to reach a decision, then there are fewer opportunities to appeal. Nevertheless, the decision to appeal more than once during the 36-month study period implies a higher level of commitment to pursuing disability benefits on the part of the participant compared to those who appeal once.

Table 4-11 provides a summary of appeal attempts. There were no significant differences in the decision to appeal or the number of appeals made by participants in each study arm. Almost half (45-46%) of participants filed at least one appeal during the study period. Among those who filed an appeal, more than half (55-60%) filed only one appeal. The average number of appeals filed among those who appealed at all was 1.7.

Table 4-11. Appeal attempts during SED study enrollment

	Measures						Regression-adjusted estimates of impact			
	Full-Service (T1) N = 976		Basic-Service (T2) N = 987		Usual Services (C) N = 981					
	n	%	n	%	n	%	T1-C	T2-C	T1-T2	AT-C
Any appeal	435	44.6	454	46.0	447	45.6	-1.62	0.49	-2.11	-0.57
Among those with any appeal										
Percent with one appeal	252	57.9	252	55.5	271	60.6	2.61	4.77	-2.16	3.72
Percent with two or more appeals	183	42.1	202	44.5	176	39.4	-1.62	0.49	-2.11	-0.57
Mean number of appeal attempts	435	1.7	454	1.7	447	1.6	0.03	0.05	-0.02	0.04

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

SSA Benefit Payments

Table 4-12 summarizes the average disability payments and average number of days until award among those who received an approval during their enrollment in the SED. Among those approved for disability, the average total disability payments made during the study was approximately $30,000 per approved participant. The treatment did not have a significant impact on these average amounts.

Table 4-12. Total disability payments paid during study period and time until award

	Averages			Regression-adjusted estimates of impact			
	Full-Service (T1) N = 143	Basic-Service (T2) N = 144	Usual Services (C) N = 152	T1-C	T2-C	T1-T2	AT-C
Average disability payments among those approved ($)	$27,399	$30,749	$30,152	-$1,873	$1,910	-$3,783	-$77
Time until award among those approved (days)	564	513	542	13.32	-28.58	41.91	-7.32

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

4.3.3 Characteristics Associated with SSA Benefit Allowance

Predicting Allowance Rates

We use a logistic regression to obtain estimates of the relationship between multiple characteristics of interest and the likelihood of SSA benefit allowance. Table 4-13 shows results from the logistic regression.

The characteristics that significantly predict allowance rates are age, health status, and work history. Older study participants, participants with lower SF-12 PCS (at baseline), and people who were not working at baseline were more likely to receive an acceptance decision. The model results indicate that for each additional year of age, the likelihood of receiving an allowance rises by 0.327 percentage points, on average. Generally, these findings are in line with the SSA determination process; that is, applicants with lower health status and less recent work experience are more likely to receive an acceptance. Furthermore, participants with higher levels of formal education (e.g., associate or bachelor's degrees) were more likely to receive an acceptance, on average, by 9.668 percentage points and 7.985 percentage points, respectively.

Better health status tends to lower the likelihood of receiving benefits. We see negative associations between benefit receipt and PCS, as well as WD-FAB domain scores including Upper Body Function and Community Mobility Drive.

Local area economic and health data also had a significant relationship with acceptance rates. Those living in states with higher levels of hospitalizations due to COVID-19, for example, were more likely to receive an acceptance to the disability rolls. It is important to note, however, that these findings as they relate to the SED study population may not generalize to the nation.

Table 4-13. Logistic regression for SSA benefit allowance (all participants)

Variable	Marginal effect[a]	Standard error	p-value
Study arm			
Reference group: Usual Services	–	–	–
Basic-Service dummy	-0.911	1.561	0.559
Full-Service dummy	-0.894	1.575	0.570
Work status at baseline			
Reference group: currently working	–	–	–
Not currently working but worked in prior 2 years	3.694	1.722	0.032
Not currently working but worked prior to the past 2 years	4.490	1.883	0.017
Never worked	0.065	5.129	0.990
Missing	10.866	11.698	0.353
Age	0.327	0.088	< 0.001
Female	-2.385	1.373	0.082
Race and ethnicity			
Reference group: White non-Hispanic	–	–	–
Black non-Hispanic	-1.425	1.638	0.384
Hispanic	-2.722	2.090	0.193
Two or more races non-Hispanic	-3.072	2.282	0.178
Other or missing	0.972	4.080	0.812

Table 4-13. Logistic regression for SSA benefit allowance (all participants) (continued)

Variable	Marginal effect^{a}	Standard error	p-value
Education			
Reference group: less than high school	–	–	–
Completed high school	3.175	1.784	0.075
Some college or technical school	3.887	1.778	0.029
Associate degree	9.668	3.049	0.002
Bachelor’s degree or higher	7.985	2.745	0.004
Nights spent in jail prior to enrollment	0.006	0.023	0.776
MCS at baseline	0.006	0.082	0.940
PCS at baseline	-0.208	0.082	0.011
Urban	2.013	2.187	0.357
FAB Domains			
Basic Mobility	-0.048	0.209	0.817
Upper Body Function	-0.366	0.208	0.078
Communication and Cognition	-0.066	0.137	0.629
Resilience	-0.050	0.090	0.576
Mood and Emotions	0.022	0.085	0.800
Interpersonal Interactions	-0.096	0.105	0.360
Community Mobility Drive	-1.260	0.580	0.030
County Average Weekly Wage – Private (BLS)	0.003	0.002	0.134
State Average Hospitalization Rate Due to COVID (CDC)	0.505	0.227	0.026
County Average Unemployment Rate (LAUS)	-0.800	0.646	0.215

Note: ^{a} Marginal effects represent average percentage point changes in the likelihood of employment during the study period. Pseudo R-squared = 0.06.

Treatment Effects on SSA Benefit Receipt by Subgroup

Although the Full-Service and Basic-Service interventions did not have significant impacts on SSA allowance rates among all participants, it is possible that the interventions did impact allowance rates among certain subgroups of participants. In Table 4-14, we provide comparisons by subgroup of interest. For each subgroup, we calculate the allowance rates for each study arm and the comparisons between each arm.

As Table 4-14 shows, we tested the impacts of the intervention within subgroups defined by age, gender, race and ethnicity, claim type, urban/rural location, education, and work history. For almost all subgroups, we found no significant impacts of the Full-Service or the Basic-Service interventions on allowance rates. There were a few notable exceptions. Men in the Full-Service group had lower allowance rates than men in the Basic-Service group, but for women we found the reverse (women assigned to Full-Service had higher allowance rates than women in Basic-Service). We also found different treatment effects depending on the participant’s reported ethnicity. Hispanic participants in the treatment arms were more likely to receive a disability approval than the Hispanic Usual Services participants.

Table 4-14. Allowance rates by demographics, claim type, work history, and education

	Percentages			Regression-adjusted estimates of impact			
	Full-Service (T1) N = 976	Basic-Service (T2) N = 987	Usual Services (C) N = 981				
	%	%	%	T1-C	T2-C	T1-T2	AT-C
Age							
Age 18-34	10.22	10.19	11.35	-0.82	-0.93	0.11	-0.87
Age 35+	17.88	17.88	18.52	-0.86	-0.78	-0.08	-0.82
Gender							
Male	12.59	17.48	14.80	-2.74	1.94	-4.68**	-0.32
Female	16.11	12.37	16.07	0.42	-3.17	3.60*	-1.36
Race and ethnicity							
White non-Hispanic	16.50	14.55	17.57	-0.46	-2.72	2.26	-1.51
Black non-Hispanic	10.70	14.97	13.55	-4.34	-0.14	-4.20	-1.96
Hispanic	15.79	14.60	12.91	7.60*	5.52	2.08	6.45*
Two or more races non-Hispanic	12.09	13.89	13.01	-1.62	0.89	-2.51	-0.56
Other/missing	20.00	12.50	26.67	-8.22	-13.68	5.46	-10.82
Claim type							
SSDI	15.82	15.42	16.69	-0.93	-0.95	0.02	-0.94
SSI	12.50	13.13	13.29	-0.82	-0.84	0.02	-0.83
Location							
Urban	15.31	14.81	15.24	-0.04	-0.43	0.39	-0.24
Rural	10.24	12.50	17.19	-6.60	-4.15	-2.45	-5.56
Education							
Less than high school	10.15	12.00	10.80	-2.12	-0.03	-2.09	-1.14
Completed high school	14.29	12.99	14.95	-0.35	-1.76	1.41	-1.10
Some college or technical	14.89	16.03	14.84	-0.16	1.76	-1.92	0.80
Associate degree	16.42	21.43	26.09	-9.51	-6.51	-2.99	-7.92
Bachelor's or higher	22.33	14.29	20.41	3.16	-5.45	8.61	-0.93
Work history at baseline							
Currently working	12.72	8.54	10.58	1.81	-1.88	3.70	-0.12
Not currently working but worked in prior 2 years	13.83	15.54	15.12	-1.89	-0.29	-1.60	-1.07
Not currently working but worked prior to the past 2 years	17.68	17.35	19.21	-0.66	-0.94	0.28	-0.79
Never worked	6.06	4.76	10.71	-5.99	-7.51	1.51	-6.70
On probation in year prior to enrollment							
Yes	10.74	12.77	14.62	-4.06	-0.57	-3.48	-2.25
No	15.24	14.96	15.65	-0.42	-0.95	0.53	-0.68

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

4.4 Health and Quality of Life

The study team hypothesized that the SED intervention would have a positive impact on health and quality life. To test this hypothesis, the study collected health and quality-of-life information at baseline and annually. We then compared these annual measures relative to the baseline measures and compared these changes over time by study arm.

4.4.1 Clinical Recovery: Mental Health

Did the SED have a significant impact on participant's mental health? On average, study participants showed significant improvement in mental health each year of the study. However, when comparing the study arms, the improvements in mental health status were not significantly larger for Full-Service or Basic-Service participants compared to Usual Services.

Table 4-15 shows the average differences in Colorado Symptom Index (CSI) scores and MCS scores, comparing the baseline score to scores measured at each year in the study (Years 1, 2, and 3). CSI scores use a 14-item assessment to assign a score on a scale from 0 to 56. The scale measures mental health symptoms: lower scores indicate fewer symptoms (better mental health). Compared to the baseline averages, each year shows a larger decline in CSI scores, suggesting that for the average participant, symptomology is decreasing (i.e., improving) over time. Comparing Year 1 CSI scores to baseline shows average declines of around 3 points. By Year 3, the difference compared to baseline was an average of around 6 points.

The team calculated MCS using items from the SF-12 questionnaire included in each annual survey. The MCS is a norm-based scale with an estimated mean for the U.S. population of 50 and a population standard deviation of 10. Scores below 50 indicate below-average mental health compared to the general population. Among each of the study arms, we see positive differences between baseline MCS and annual scores. By the end of the third year in the study, participants in each study arm saw an average improvement of just over 5 points in MCS scores, representing approximately a one-half standard deviation increase. Comparing the study arms, however, shows no significant differences in improvements in MCS from the Full-Service or Basic-Service interventions compared to the Usual Services group.

Table 4-15. Difference in SF-12 MCS and CSI at baseline and study exit

Variable	Weighted measures						Regression-adjusted estimates			
	Full-Service (T1)		Basic-Service (T2)		Usual Services (C)					
	n	#	n	#	n	#	T1-C	T2-C	T1-T2	AT-C
Colorado Symptom Index (CSI)										
Year 1 difference from baseline	692	-3.12	724	-2.74	664	-3.32	0.26	0.37	-0.10	0.32
Year 2 difference from baseline	675	-4.87	707	-4.48	659	-5.61	0.77	0.86	-0.09	0.82
Year 3 difference from baseline	623	-6.62	641	-5.54	570	-6.43	-0.27	0.57	-0.84	0.16
SF-12 MCS scores										
Year 1 difference from baseline	669	3.08	681	2.30	646	3.24	-0.33	-0.56	0.22	-0.45
Year 2 difference from baseline	645	4.83	670	4.19	640	5.21	-0.68	-0.89	0.21	-0.79
Year 3 difference from baseline	604	5.61	610	5.04	556	5.67	-0.38	-0.28	-0.09	-0.33

Notes: *** $p < .01$, ** $p < .05$, * $p < .10$. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

Characteristics Associated with Changes in Mental Health Status

We explored characteristics that may associate with changes in mental health status. Table 4-16 shows the results of the covariates included in the regression for the difference between Year 3 MCS and baseline MCS. Of the characteristics tested, a small number showed a significant relationship with changes in MCS. Older participants, on average, showed smaller improvements in MCS. Women also had smaller improvements in MCS on average compared to men. Compared to other groups defined by race and ethnicity, Hispanic participants had smaller improvements in MCS, on average. Higher baseline MCS were associated with smaller improvements in mental health status over the 3-year study period. This finding stands to reason as participants with higher initial MCS had less room for improvement compared to others.

Four of the WD-FAB domain measures (measured at baseline) had a significant relationship with changes in MCS over the study enrollment period. Resilience, Mood and Emotions, Interpersonal Interactions, and Community Mobility Drive all had significant positive relationships with changes in MCS.

Table 4-16. Linear regression of change from baseline to study exit in mental health status (SF-12 MCS)

Variable	Coefficient	Standard error	p-value
Study arm			
Reference group: Usual Services	–	–	–
Basic-Service dummy	-0.28	0.67	0.67
Full-Service dummy	-0.38	0.69	0.58
Work status at baseline			
Reference group: currently working	–	–	–
Not currently working but worked in prior 2 years	0.71	0.81	0.38
Not currently working but worked prior to the past 2 years	-0.33	1.13	0.77
Never worked	-0.45	2.04	0.83
Missing	6.68	6.70	0.32
Weeks employed (past 2 years baseline)	0.02	0.01	0.22
Total earnings (past 2 years baseline)	0.00	0.00	0.37
Age	-0.07	0.04	0.05
Female	-1.20	0.61	0.05
Race and ethnicity			
Reference group: White non-Hispanic	–	–	–
Black non-Hispanic	-0.02	0.69	0.98
Hispanic	-1.98	0.93	0.03
Two or more races non-Hispanic	0.37	1.12	0.74
Other or missing	-0.88	1.78	0.62
Education			
Reference group: less than high school	–	–	–
Completed high school	-0.05	0.85	0.95
Some college or technical school	-0.85	0.84	0.31
Associate degree	1.46	1.27	0.25
Bachelor’s degree or better	0.02	1.13	0.99
Nights spent in jail prior to enrollment	0.00	0.01	0.96
SNAP receipt in 12 months prior to enrollment	-1.02	0.65	0.12
TANF receipt in 12 months prior to enrollment	0.38	1.04	0.71
MCS at baseline	-0.71	0.04	0.00
PCS at baseline	-0.01	0.04	0.75
Urban	0.02	0.90	0.98
FAB Domains			
Basic Mobility	-0.08	0.09	0.35
Upper Body Function	-0.01	0.09	0.88
Communication and Cognition	-0.06	0.06	0.28
Resilience	0.10	0.04	0.01
Mood and Emotions	0.20	0.04	0.00
Interpersonal Interactions	0.07	0.04	0.09
Community Mobility Drive	0.56	0.30	0.07
County Average Weekly Wage – Private (BLS)	0.00	0.00	0.64
State Average Hospitalization Rate Due to COVID (CDC)	0.13	0.10	0.20
County Average Unemployment Rate (LAUS)	0.34	0.27	0.21

Notes: Weighted ordinary least squares (OLS) regression coefficients; N = 1,770. R-squared = 0.30.

Clinical Recovery: Physical Health

From the SF-12 survey items, the team also calculated the PCS at baseline and at the end of each year of study enrollment. Table 4-17 provides a summary of the differences between each annual PCS and the baseline score taken at the time of study enrollment. The positive weighted measures indicate that each study arm on average showed improvements in PCS compared to baseline. Compared to the baseline PCS, the Year 2 scores show significant positive improvements among Full-Service and Basic-Service participants relative to Usual Services participants. The magnitude of the differences is between 1 and 2 scale points, on average. In the third year of study enrollment, we see larger gains in PCS for the Full-Service relative to the Basic-Service participants.

Table 4-17. Differences in SF-12 PCS between annual surveys and baseline

Variable	Weighted measures						Regression-adjusted estimates			
	Full-service (T1)		Basic-service (T2)		Usual services (C)					
	n	#	n	#	n	#	T1-C	T2-C	T1-T2	AT-C
Physical health (SF-12 PCS)										
Year 1 difference from baseline	669	1.29	681	1.73	646	1.09	0.48	0.68	-0.20	0.58
Year 2 difference from baseline	645	2.48	670	2.01	640	1.15	1.73***	1.13**	0.60	1.43***
Year 3 difference from baseline	604	2.57	610	1.51	556	2.12	0.93	-0.58	1.51***	0.17

Notes: *** p < .01, ** p < .05. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

Quality of Life

We record two quality-of-life measures using the annual surveys: satisfaction with life, and number of arrests.

Satisfaction with life is the difference in the participant's scores from the Quality of Life scale, measured annually. The Quality of Life scale is a 7-point measure ranging from Terrible (1) to Delighted (7) in response to the question "how do you feel about your life in general?" The average baseline score for all study participants was 3.49, which translates to a typical response between "mostly dissatisfied (3)" and "mixed (4)." The average increases over the course of the study are generally around a half point on the scale. Table 4-18 shows the differences in quality-of-life scores over the study period for each study arm. When comparing the study arms, we see no significant differences in the changes in quality-of-life scores at each annual survey over the course of the study.

Table 4-18. Difference in quality-of-life scores between annual surveys and baseline

Variable	Weighted measures						Regression-adjusted estimates of difference-in-difference			
	Full-Service (T1)		Basic-Service (T2)		Usual Services (C)					
	n	#	n	#	n	#	T1-C	T2-C	T1-T2	AT-C
Satisfaction with life (score)										
Year 1 Dif.	684	0.46	719	0.43	659	0.55	-0.07	-0.08	0.01	-0.07
Year 2 Dif.	672	0.65	704	0.67	654	0.77	-0.12	-0.08	-0.04	-0.10
Year 3 Dif.	620	0.89	634	0.83	569	0.88	0.01	0.01	0.00	0.01

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

The study measured criminal justice involvement by asking participants to indicate on each annual survey whether they were arrested in the past 12 months and the number of arrests. Table 4-19 provides a summary of the average number of arrests per study arm at each of the three annual surveys taken throughout enrollment in the study. The participants in each study arm are not statistically different from one another in terms of number of arrests, indicating that, by this measure, the SED Full-Service and Basic-Service interventions did not have a significant impact on criminal justice involvement.

Table 4-19. Number of arrests by study arm

Variable	Weighted measures						Regression-adjusted estimates of impact			
	Full-Service (T1)		Basic-Service (T2)		Usual Services (C)					
	n	#	n	#	n	#	T1-C	T2-C	T1-T2	AT-C
Number of Arrests										
Year 1	685	0.19	716	0.12	659	0.12	-0.01	-0.06**	0.05	-0.04
Year 2	667	0.10	698	0.09	652	0.10	0.06	-0.01	0.07	0.03
Year 3	620	0.10	633	0.08	564	0.09	0.01	<0.01	0.01	<0.01

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

4.5 Utilization of Services

The study team collected information on hospital and outpatient visits throughout the study. Participants reported each ER visit, hospital overnight stay, and outpatient hospital visit. We hypothesized that the interventions would reduce reliance on healthcare services. To test this hypothesis, we compared the number of encounters indicated by participants assigned to each study arm. Table 4-20 summarizes overall service use for each category over the course of the 3-year study enrollment. The measures of healthcare service are:

- **ER Visits.** These are the sum of all reported visits to an ER. For instructional purposes, we also report a breakdown of ER visits for a (1) mental health problem and (2) physical health problem. On average, participants in each study arm had just over two ER visits during the 3-

year study period due to a physical health problem, and less than one visit due to a mental health problem.

- **Hospital Overnight Stays.** These are the sum of all reported overnight stays in a hospital. For instructional purposes, we also report a breakdown of overnight stays for (1) a mental health problem (including drug or alcohol problem) and (2) a general health problem. The average number of hospital overnight stays during the 3-year study period was just under one visit.

- **Outpatient Hospital Visits.** Outpatient hospital visits are the sum of all medical visits (surgeries or other procedures) to a hospital not resulting in an overnight stay. For instructional purposes, we also report a breakdown of outpatient visits by (1) preventive care, (2) physical health, and (3) other. The average number of outpatient visits was just under one visit.

Table 4-20. Service utilization measures by study arm

Variable	Weighted measures						Regression-adjusted estimates of impact			
	Full-Service (T1)		Basic-Service (T2)		Usual Services (C)					
	n	#	n	#	n	#	T1-C	T2-C	T1-T2	AT-C
ER visits	582	2.38	599	2.36	541	2.59	-0.28	-0.16	-0.12	-0.22
For mental health problem	582	0.26	599	0.31	541	0.25	-0.02	0.05	-0.07	0.02
For physical health problem	582	2.15	599	2.11	541	2.44	-0.32	-0.24	-0.07	-0.28
Number of nights spent in hospital	582	3.05	599	2.67	541	3.30	-0.32	-0.37	0.04	-0.34
Hospital overnight stays										
After ER visit for physical problem	582	0.47	599	0.45	541	0.55	-0.05	-0.06	0.01	-0.06
After ER visit for mental problem	582	0.14	599	0.19	541	0.12	0.01	0.07	-0.06*	0.04
Hospital stay for physical problem	582	0.19	599	0.17	541	0.17	0.01	0.01	0.01	0.01
Hospital stay for mental problem	582	0.04	599	0.07	541	0.06	-0.03*	0.00	-0.03**	-0.01
Outpatient hospital visit/ procedure	582	0.87	599	0.78	541	0.78	0.08	0.03	0.06	0.06
Preventive care	582	0.20	599	0.21	541	0.15	0.05	0.07*	-0.02	0.06**
Physical health	582	0.74	599	0.59	541	0.66	0.07	-0.06	0.13*	0.01
Other problem	582	0.17	599	0.14	541	0.16	0.01	-0.02	0.03	-0.01

Notes: ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

Most utilization measures show no significant impacts of the SED Full-Service or Basic-Service interventions. However, there are some notable exceptions.

Hospital Stays and ER Visits for Mental Problems. Participants in the Full-Service group had slightly fewer hospital stays and ER visits for mental problems compared to the Basic-Service group.

Preventive Care. The SED had a marginal impact on the likelihood of completing an outpatient visit for preventive care. Participants in the Full-Service and Basic-Service groups were each more likely than the Usual Services participants to complete preventive care visits during their enrollment in the study.

The annual surveys asked participants to report the number of routine visits (non-hospital and non-ER) they completed during the past month prior to the survey. Table 4-21 provides a summary of the average number of routine visits at baseline and at each annual survey. The types of visits summarized in the table are:

- **Routine mental health visits.** Defined as the number of non-hospital or non-ER visits for a mental health problem (includes drug and alcohol problems and visit to a psychiatrist, other mental health professional, and self-help groups) in the past month. We also report routine mental health visits in the past month for each year of study participation and for each quarter of study participation.

- **Routine general health visits.** Defined as the number of non-hospital and non-ER visits for a general health problem in the past month. We also report routine general health visits in the past month for each year of study participation and for each quarter of study participation.

- **Routine employment support visits.** Defined as the number of visits in the past month for vocational or educational training, finding a job, or keeping a job. We also report routine employment support visits in the past month for each year of study participation and for each quarter of study participation.

We see no significant impacts of the Full-Service or the Basic-Service interventions on the numbers of routine mental health or general health visits at each year. We do, however, see differences in the numbers of routine employment support visits. In each year, the Full-Service and Basic-Service participants had more employment support visits than the Usual Services, on average. This result is expected given that both the Full-Service and Basic-Service groups received IPS SE as a core part of both treatment interventions.

Table 4-21. Routine outpatient service utilization measures by study arm

Variable	Weighted measures						Regression-adjusted estimates of impact			
	Full-Service (T1)		Basic-Service (T2)		Usual Services (C)					
	n	#	n	#	n	#	T1-C	T2-C	T1-T2	AT-C
Routine Mental Health Visits										
Baseline	976	2.22	987	2.11	980	2.28	-0.08	-0.13	0.05	-0.10
Year 1	692	1.81	724	1.63	663	1.49	0.24	0.18	0.06	0.21
Year 2	676	1.32	707	1.46	658	1.26	0.00	0.19	-0.19	0.10
Year 3	622	1.32	641	1.16	569	1.31	-0.10	-0.19	0.09	-0.14
Routine General Health Visits										
Baseline	976	1.08	987	1.07	981	1.09	-0.04	-0.02	-0.03	-0.03
Year 1	692	0.84	723	0.80	664	0.77	0.03	0.04	-0.01	0.03
Year 2	676	0.66	706	0.67	658	0.62	0.00	0.05	-0.04	0.03
Year 3	622	0.73	640	0.63	570	0.65	0.07	0.01	0.06	0.04
Routine Employment Support Visits										
Baseline	973	0.36	987	0.34	978	0.35	0.01	-0.04	0.05	-0.01
Year 1	690	0.74	721	0.69	662	0.35	0.40***	0.36***	0.04	0.38***
Year 2	672	0.43	703	0.55	655	0.23	0.19**	0.31***	-0.13*	0.25***
Year 3	624	0.42	639	0.44	567	0.22	0.19**	0.21**	-0.02	0.20***

Notes: *** p < .01, ** p < .05, * p < .10. Impact estimates are regression-adjusted. Weighted percentages do not necessarily equal weighted counts. Regression-adjusted estimates of impact may not be consistent with differences in weighted means. AT = All Treatment (T1 and T2 participants combined).

5. Cost-Benefit Analysis

5.1 Introduction

Key Findings

- During the 3-year study period, the total program costs outweigh the participant benefits (earnings and fringe employment benefits) for both the Full-Service and Basic-Service groups.
- The benefits increased sharply in Year 3. These trends and persuasive evidence from other studies suggest these 3-year net benefit comparisons grossly understate net benefits over a somewhat longer timeframe than the 36-month study period.
- Relative to the Usual Services group, treatment group participants saw the largest earnings benefits from SED in their third year of study enrollment. It is possible that higher earnings among treatment group participants may continue beyond the study period.
- SED labor costs may overestimate costs in a typical IPS setting due to the payment structure.

The cost-benefit analysis (CBA) assesses whether the benefits of the SED interventions outweigh the costs associated with the interventions. The results of this analysis will help to address the critical question of whether either or both types of SED interventions (Basic-Service and Full-Service) merit further investment. This chapter presents findings of the SED CBA. We discuss the accounting framework that was used; procedures to monetize benefits and costs; the steps necessary to determine the present value of net benefits; as well as plans for extrapolating these estimates to future years.

5.2 Accounting Approach

The basic approach to assessing the economic efficiency of using scarce resources to implement any public policy is simple in concept. Estimate and evaluate the effects of the policy (i.e., the difference between having and not having the policy) on the well-being of all affected persons.

Policy effects that increase (decrease) an individual's well-being (based on their own preferences) are benefits (costs) to that individual. For some individuals, the combination of effects is net positive while for others the combination is net negative. According to the basic logic of BCA, if the sum of net positive effects for the former persons (i.e., the "winners") exceeds the sum of net negative effects for the latter group (i.e., the "losers"), the overall sum of net benefits for society as a whole is positive and the policy should be adopted. This is the fundamental rule of policy analysis.

Note that in most (if not all) policy interventions, the same person can experience both benefits and costs. For example, in the SED many participants experienced benefits deriving from increased employment and earnings as well as costs since they had to pay some of the earnings as government taxes. Thus, while BCA analysis distinguishes among specific positive and negative effects, and it can provide net benefit assessments for subgroups (e.g., participants) as well as for all affected persons (i.e., "society"), it does not presume that affected persons can only experience one type of effect. Thus, it does not define net benefits from the "perspective" of taxpayers. Similarly, it does not measure "net benefits" for non-persons such as governments, government agencies, or organizations. As explained later, however, there often are good reasons to examine net budgetary effects on governments of a policy as an intermediate step in the BCA evaluation process.

5.2.1 Costs of the Intervention

The costs of providing SED services fall into two broad categories: (1) labor and overhead costs, and (2) other behavioral health and work-related expenses. The study also provided compensation for necessary healthcare for uninsured participants. Only the participants randomized into the Full-Service or the Basic-Service study arms received SED intervention services. All uninsured SED participants (including Usual Services participants) had access to healthcare compensation.

Labor and Overhead Costs. Sites received per-participant payments calculated based on the number of SED participants enrolled each month at the site. The team determined labor costs based on the applicable local labor market rates for each SED treatment team member. Sites then received payments for labor and overhead based on the numbers of Full-Service and Basic-Service participants enrolled at the site.

Labor costs included wages for SED team members: care managers, IPS specialists, IPS supervisors, and NCCs. Both Basic-Service and Full-Service participants received services from IPS supervisors, IPS specialists, and care managers. Only Full-Service participants received services from NCCs; therefore, the difference in labor costs between Basic-Service and Full-Service participants are due entirely to labor costs for NCCs.

Table 5-1 shows the expected number of annual hours per position for a large site (assuming 40 enrollees each for Basic-Service and Full-Service) and a small site (assuming 20 enrollees each for Basic-Service and Full-Service).

Table 5-1. Budgeted hours for 1 year of basic-service and Full-Service labor and overhead

Position	Small sites		Large sites	
	Basic-Service	Full-Service	Basic-Service	Full-Service
Case manager	696	696	1,392	1,392
IPS specialist	1,685	1,685	3,370	3,370
IPS supervisor	246	246	492	492
NCC	0	1,030	0	2,059

Notes: Budgeted hour totals are for 1 year of services with full enrollment for the entire year. Assumes 80 participants (40 Full-Service and 40 Basic-Service) for large sites and 40 participants (20 Full-Service and 20 Basic-Service) for small sites.

Startup Costs. Each site received a payment of $13,000 when signing the contract to participate in the SED. This payment covered initial startup costs for the site. Sites then received monthly payments designed to cover staff hours, training, and overhead for the first 14 months while the study team recruited participants. For the next 14 months, sites then received equal payments based on the expected number of study enrollees for the site (40 per study arm for large sites and 20 per study arm for small sites). The SED team structured these payments in equal installments through these initial 14 months to provide consistent funding to hire and train staff as new participants enrolled during the recruitment period.

Benefits Training. The SED provided benefits planning training to an identified staff member from each demonstration site to ensure that each site had at least one Certified Work Incentives Counselor (CWIC) who could provide formal benefits planning to SED participants with SSA benefits, in addition to any other benefits counseling SED participants may have received. Under subcontract to Westat, the Virginia Commonwealth University National Training and Data Center provided to trainees a 5-day in-person initial training, certification assessments, ongoing technical assistance as needed, and continuing certification activities.

Average Intervention Costs Per Participant

Table 5-2 provides a summary of average per-participant labor costs, behavioral health, and work-related expenses borne by SSA for the 36-month study period, separately for Full-Service and Basic-Service participants. Across all sites combined, the average per-participant labor and overhead cost, including clinical technical assistance, for Full-Service participants was $22,184, and the average Basic-Service cost was $13,343.

Table 5-2. Average per-participant intervention costs by study arm over 36-month study period paid for by SSA

Cost type	Full-Service	Basic-Service
Labor costs		
Labor and overhead costs		
Startup costs (initial fixed payments)	$196	$194
Labor costs	$21,251	$12,543
Clinical Technical Assistance	$504	$375
Benefits training	$233	$231
Total labor costs	**$22,184**	**$13,343**
Other behavioral health and work-related expenses		
Clinical behavioral health	$278	$159
Clinical general medical[a]	$110	$93
Behavioral health medications	$56	$30
General health medications	$20	$16
Individual work-related expenses	$143	$164
Nonclinical support	$708	$709
Uninsured claims		
Uninsured claims	$197	$142
Administration costs[b]	$60	$60
Total other expenses and uninsured claims	$1,572	$1,373
Total costs	**$23,756**	**$14,716**

Note: Averages are across all participants for the 36-month study period.

[a] Excludes expenses for hospital inpatient admissions, ER visits, and outpatient visits to avoid double-counting healthcare utilization costs measured by the quarterly surveys.

[b] Administrative costs cover staff time needed to review and approve payment claims.

It is important to note that the average per-person labor and overhead costs depend on the total number of enrollees randomized to the site. The sites continued to receive payments for providing services to all treatment participants regardless of whether these participants took advantage of the services offered, or whether the participant died or withdrew from the study.

In addition to labor costs for providing services, the average labor and overhead costs include the initial startup costs paid to the sites prior to enrollment of study participants. These costs covered hiring, onboarding, and training staff.

The 30 SED sites recruited staff from diverse labor markets with different prevailing local labor rates for IPS specialists, IPS team leads, NCCs, and care managers. The study team assigned hourly rates for each staff position by site and year. The hourly rates included wages, fringe benefits, and overhead. The SED team determined hourly rates for each site based on competitive rates in each site's local labor market. This meant that the cost of staffing SED sites varied across sites. The average site-specific cost for Full-Service participants ranged from $14,760 to $28,781 per participant; for Basic-Service participants, the average per-participant costs ranged from $9,541 to $17,087.

Clinical Technical Assistance. The implementation team supported sites through clinical consultations and intensive case reviews. This support was available to each site for Full-Service and Basic-Service participants. To estimate the cost of providing this assistance, the evaluation team estimated hours spent providing support and multiplied these hours by the hourly rates of the implementation team members. Costs for all but one implementation team member were available to both Full-Service and Basic-Service participants; we divided labor costs for these team members evenly among the Full-Service and Basic-Service participants. The remaining team member supported only the NCCs; therefore, we applied the labor costs for this team member only to the Full-Service participant averages.

Other Behavioral Health and Work-Related Expenses. The SED provided funding for other costs associated with behavioral health and work. These reimbursements covered costs that the team organized into general categories based on the type of need addressed. The averages in Table 5-2 include all eligible Full-Service and Basic-Service enrollees, including those who did not take advantage of services through the site or through reimbursements.

- **Clinical Behavioral Health Services.** These services include psychotherapy, individual, or family counseling, physical healthcare, psychiatric consultation, physical therapy, and occupational therapy. Clinical services for behavioral health averaged $278 per Full-Service participant and $159 per Basic-Service participant.

- **Clinical General Medical.** These included payments for visiting a clinic for checkups or other minor health problems, outpatient visits, dental exams and procedures, as well as medical equipment necessary for work such as glasses. General medical expenses averaged $110 per Full-Service participant and $93 per Basic-Service participant.

- **Behavioral Health–Related Medication Expense.** Coverage includes deductibles, co-pays, and full cost of prescription drugs for the treatment of mental health symptoms, if not covered by the participant's health insurance. These costs averaged $56 per Full-Service participant and $30 per Basic-Service participant for behavioral medication costs, and $20 and $16, respectively, for general health medications for Full-Service and Basic-Service participants.

- **Individual and Work-Related Expenses (IWRE).** These consist of items or services directly related to taking a specific job and are typically associated with IPS service delivery. Examples include business-appropriate attire, certifications, licensures, and transportation costs for interviews. In special situations, it would also include dentures or other dental services that may alleviate a barrier to entering a job. These costs averaged $143 per Full-Service participant and $164 per Basic-Service participant.

- **Non-clinical support services.** These consist of other items needed to help participants overcome barriers impeding their return to work. Typically associated with care management, these include services and expenses for temporary, short-term, or emergency assistance to address housing, legal, or transportation barriers. The reimbursement must enable the participant to overcome the barrier completely and not represent an ongoing need. Nonclinical support services comprised the largest category of other behavioral health and work-related expenses, averaging $1,557 for Full-Service participants and $1,344 for Basic-Service participants. Most nonclinical support service reimbursements (53%) covered housing costs (e.g., mortgage, rent, shelter costs).[4]

Uninsured Claims. In addition to behavioral health and work-related expenses, the SED study paid for necessary healthcare expenses for study participants who did not have access to healthcare coverage. On a per-participant basis, the average cost of covering uninsured claims was $197 for Full-Service participants and $142 for Basic-Service participants. When budgeting for the study, the team assumed that approximately 10 percent of enrollees would not have access to healthcare. The actual number of enrollees without healthcare coverage at the time of enrollment was closer to 20 percent.

5.2.2 Calculating Benefits

This section monetizes impacts from two types of benefits: benefits from work, and reductions in healthcare utilization costs. We divide benefits from work into two types: earnings and fringe benefits. Table 5-3 shows the average per-participant benefits from work of the interventions.

Earnings. Participants in the treatment groups had significantly higher employment rates and earnings than participants in the Usual Services group. The benefits from each intervention (Full-Service and Basic-Service) are the average per-participant increases over the Usual Services group.

Fringe Benefits. The quarterly follow-up survey asked SED participants to indicate the benefits available to them from any jobs that they worked since the previous survey. We used the responses from the quarterly surveys combined with information on employer spending on benefits from the Bureau of Labor Statistics to estimate the total fringe benefits for each participant (U.S. Bureau of Labor Statistics, 2020a and 2020b). We based calculations of benefits received on the average cost of providing benefits to employees who receive benefits as a percentage of total compensation and the total earnings at the job providing benefits as reported on the quarterly surveys. In terms of dollars, the largest category of benefits was health insurance. On average, both Full-Service and Basic-Service participants received over $250 in health insurance benefits more than the Usual Services group over the 36-month study period.

---

[4] For a detailed breakdown of costs for IWRE and nonclinical support services, see the *Final Process Analysis Report (Deliverable 7.5).*

Full-Service and Basic-Service participants received an average of $4,880 and $3,493 in earnings and fringe benefits more than the Usual Services group.

Table 5-3. Average per-participant earnings and fringe benefits of Full-Service and Basic-Service participants compared to control over 36-month study period

Benefit type	Full-Service vs. control	Basic-Service vs. control
Earnings	$4,265	$3,027
Fringe benefits from work		
Health insurance	$336	$279
Retirement	$91	$45
Vacation time	$150	$112
Sick leave	$38	$30
Total benefits from work (Earnings + Fringe benefits)	$4,880	$3,493

Notes: Cell entries are estimated treatment effects based on regressions controlling for baseline characteristics, local area economic conditions, and other factors related to employment and earnings.

Among the SED Full-Service and Basic-Service participants, earnings benefits from the study increase in each year of study enrollment. SED service utilization, measured as the average number of IPS specialist and care manager visits per year of enrollment, decrease with each enrollment year. Combined, these trends suggest, on average, that benefits continue to increase in the third year of study enrollment. As Table 5-4 shows, the largest earnings benefits from the intervention occurred in the third year of study enrollment.

Table 5-4. Average per-participant benefits of Full-Service and Basic-Service participants compared to control by study year

Benefit type	Year 1		Year 2		Year 3	
	Full-Service	Basic-Service	Full-Service	Basic-Service	Full-Service	Basic-Service
Earnings	$394	$426	$1,146	$604	$2,725	$1,978
Fringe benefits from work						
Health insurance	$20	$42	$84	$70	$233	$165
Retirement	$12	$9	$35	$10	$45	$23
Vacation time	$8	$21	$31	$13	$111	$77
Sick leave	$2	$5	$9	$5	$27	$21
Total benefits from work	$436	$503	$1,305	$702	$3,141	$2,264

Notes: Cell entries are estimated treatment effects based on regressions controlling for baseline characteristics, local area economic conditions, and other factors related to employment and earnings.

Reduction in Healthcare Service Utilization. The study team estimated healthcare utilization costs for each study arm and compared the study arms using the same regression-adjusted modeling approach used in Chapter 4. Table 5-5 includes estimates of total costs of healthcare service utilization based on survey responses and the regression-adjusted estimates of impact. We estimate these costs using responses to the annual surveys in which participants indicated the total number of nights spent in the hospital, the total number of emergency room visits, and the total number of outpatient visits made during their 3-year study enrollment. We then multiplied these visits by the average costs in 2021 dollars using national estimates obtained from the Department of Health and Human Services (Freeman, Weiss & Heslin, 2018; Moore & Liang, 2020). On average, Full-Service participants had $10,277 in combined hospital inpatient stays, ER visits, and outpatient visits during the 3-year study period, compared to $9,378 for Basic-Service and $10,873 for Usual Services participants.

Although the weighted means differ, the regression-adjusted estimates indicate that differences in weighted means for healthcare utilization costs are not statistically significant across the study arms. The lack of significance was expected given that the average number of hospital inpatient stays, ER visits, and outpatient visits did not differ significantly across the study arms.

Table 5-5. Average healthcare utilization costs by study arm over the 36-month study period

	Weighted measures			Regression-adjusted estimates of impact			
	Full-Service (T1) N = 582	Basic-Service (T2) N = 599	Usual Services (C) N = 541	T1-C	T2-C	T1-T2	AT-C
Hospital Inpatient stays	$8,728	$7,833	$9,229	-$1,158	-$1,431	$273	-$1,290
ER visits	$1,408	$1,416	$1,529	-$160	-$112	-$48	-$136
Outpatient visits/procedures	$141	$129	$115	$15	$6	$9	$11
Total	$10,277	$9,378	$10,873	-$1,307	-$1,540	$233	-$1,420

Source: Based on weighted average incidences from the SED quarterly follow-up survey. Costs of hospital inpatient stays and ER visits are based on average national costs from Health and Human Services (HHS) Agency for Healthcare Research and Quality (AHRQ), standardized to 2021 dollars. Regression-adjusted estimates of impact are estimated treatment effects based on regressions controlling for baseline characteristics, local area economic conditions, and other factors related to employment and earnings.

5.2.3 Comparing Benefits and Costs During the Study Period

Table 5-6 summarizes the total costs and estimated benefits per participant for the 3-year intervention period. The only outcome measures that showed significant difference between the participants in the randomly assigned SED study arms were employment and earnings. Therefore, the benefits estimates include only differences in earnings and fringe benefits. Reductions in healthcare utilization costs are not included because there was no significant difference between the Full-Service or Basic-Service groups and the Usual Service based on the regression-adjusted estimates of impact.

Estimating Net Impact on Service Costs

As part of the overall CBA of the SED intervention, it is important to estimate the net impact on service use costs defined as the difference in costs between the treatment groups' participants and the control (i.e., "Usual Service") group participants. For the treatment groups, some of these costs, related to services provided by the sites and other services paid for under the project, are included in the overall accounting data on SSA project expenditures. Analogous cost data for the same types of services used by control participants, however, were not gathered as part of the project. Estimated quantities of use for a number of these types of services were collected in self-reports via the baseline and quarterly follow-up interviews.

Table 5-6. Average Full-Service and Basic-Service per-participant benefits and costs for the intervention period relative to usual services

Estimated costs and benefits	Low estimate		High estimate	
	Full-Service	Basic-Service	Full-Service	Basic-Service
Total intervention costs	$23,756	$14,716	$23,756	$14,716
Usual Care uninsured costs	$98	$98	$98	$98
Estimated Usual Services employment support costs	$2,411	$2,411	$5,942	$5,942
Total intervention costs minus Usual Services costs	$21,247	$12,207	$17,716	$8,676
Total intervention benefits	$4,880	$3,493	$4,880	$3,493
Net benefit-cost	-$16,367	-$8,714	-$12,836	-$5,183

Since our impact estimates, as noted above, found no significant differences in these self-reported quantities between the control and treatment groups, and lacking information on possible differences in the per unit cost of these services, the logic of CBA suggests these service costs that are already included in the intervention expenditures by SSA should in fact be deducted from the intervention expenditure figures used on the cost side of the CBA. Thus, the empirical challenge is to estimate and separate the portion of intervention costs that should in fact be excluded from the CBA tax funds needed to pay for these costs. (We abstract here from any CBA adjustment to reflect the societal "excess burden" efficiency costs of the need to use taxes to fund these excluded intervention expenditures.)

Estimating Usual Services (Control) Group Costs

After randomization, the Usual Services group received information about mental health and employment services available in the local area. The study team did not collect information about the cost of services that the Usual Services group received. However, the quarterly survey asked all study participants, including Usual Services group participants, to indicate the number of routine visits in the month prior to the survey for employment support. We use the self-reported survey data to estimate potential costs for similar employment support services received by the control group.

To estimate costs of employment support, we assigned an average yearly cost value for employment support. Because we do not have a complete accounting of the total number of employment support visits made throughout the study period, we calculated lower and upper bound estimates of total employment support costs for the control group. As an upper bound on employment support costs for the Usual Services group, we assumed that any participant who completed at least one routine employment support visit in the month prior to completing a quarterly survey received employment support services throughout the year. For the upper bound estimate, we assigned $6,100 for each year of study enrollment to participants who indicated on the survey that they completed at least on employment support visit during the year.[5] Following this methodology, the average per-person employment support costs for the Usual Services group over the 3-year study period are $5,942. As a lower bound estimate, we divided the average yearly cost of $6,100 by four to obtain a quarterly cost ($1,525) and assumed that any participant who completed at least one employment support visit in the previous quarter received employment support for that quarter, then summed the costs for the total number of quarters of employment

---

[5] The $6,100 estimate is based on the average annual estimate of $5,000 for IPS employment support costs from 2013, adjusted for inflation to 2021 dollars. See Salkever, D.S. (2013).

support received over the 3-year study period. Using this approach, the estimated average per-participant employment support cost for the control group members is $2,411.

To ensure a level playing field for all study enrollees, Usual Services participants also had the opportunity to file claims for healthcare expenses if they did not have health insurance. A small number (n = 11) of Usual Service participants made uninsured claims. The cost of paying uninsured claims as a per-participant average across all Usual Service participants was $38, and the administrative cost was $60, for a total of $98.[6] To estimate net costs and benefits, we subtract these costs from the total intervention costs for the treatment groups.

Looking at only benefits realized during the 3-year study period, the total costs outweigh the benefits for both the Full-Service and Basic-Service groups. For the low estimate, we use the lower bound estimate for employment support costs for the Usual Services group ($2,411). For the low estimate for the average Full-Service participant, the total costs outweighed benefits by $16,367, and for the average Basic-Service participant, total costs outweighed estimated benefits by $8,714. For the high estimate, we use the upper bound estimate of employment support costs for the Usual Services group ($5,942). For the high estimate for the average Full-Service participant, the total costs outweighed benefits by $12,836, and for the average Basic-Service participant, total costs outweighed estimated benefits by $5,183.

5.3 Discussion

Although the benefits did not outweigh costs during the intervention period for either the Full-Service or Basic-Services participants, the net benefit figures presented above substantially understate the net benefits of the intervention when a slightly longer timeframe is employed. Similarly, there is good reason to believe that this understatement of net benefits applies to future replications of the SED intervention.

There is evidence from follow-up studies of IPS interventions that benefits from work extend beyond the initial period of intervention (Becker, Whiteley, Bailey & Drake, 2007; McHugo, Drake & Becker, 1998). However, whether these benefits continue for individuals who no longer receive SED services is an open question that merits further investigation.

There is evidence that suggests that benefits from work may continue for treatment group participants even after study enrollment ends. As discussed in the SED *Final Process Evaluation Report*, engagement in SED services (as measured by total meetings with site staff), are on average highest in the first year of study enrollment and decrease in Years 2 and 3. Combined with the upward trend in earnings benefits during study enrollment, these trends suggest that the benefits are sustained by the average treatment group participant despite lower levels of engagement with SED supports.

Furthermore, we see that later enrollees received greater benefit from participation in the SED treatment than earlier enrollees. This finding suggests that fully established programs may generate greater economic benefits than those in the startup phase who are still learning to serve denied disability applicants. Therefore, established programs that have experience serving this population may be more cost effective.

---

[6] Because the same staff member administered the uninsurance claims for all SED participants, administrative costs were divided evenly among the three study arms.

5.4 Limitations

There are notable uncertainties around estimates of the costs and benefits. This section briefly mentions some of the limitations associated with this CBA.

A major limitation of the BCA and budgetary impact calculations discussed previously is that they are limited to the 3-year duration of the intervention and the associated evaluation. There is reason to believe that limiting the period of the evaluation to exclude any post-intervention experience may substantially understate the net benefits of the SED since employment and earnings gains from the SED will be ignored while the post-intervention costs will be essentially zero. Experience from longer-term follow-up studies of IPS SE participants with mental disorders suggests that ongoing earnings gains could be significant.

Another limitation to note is that the costs presented here may overestimate service costs for the typical IPS recipient. The study funded labor costs to sites based on their total enrollment of SED treatment participants, regardless of whether participants used services or formally withdrew from the study. SED site staff continued to try to engage SED participants throughout the 3-year study period even if the participant never engaged with the site. In a typical IPS setting, staff would not continue to attempt to contact participants who did not engage with a site for over a month. In the SED, however, sites continued to receive payments to serve participants who did not engage with the site even though these participants stood little chance to benefit from their enrollment in the study.

6. Synthesis of Process Evaluation and Quantitative Findings on Outcomes

6.1 Introduction

This chapter synthesizes the findings from the process evaluation and the quantitative outcomes of the SED. We focus on the main findings of the impact analyses and provide contextual information to discuss potential explanations for each finding.

6.2 Participants in Both Treatment Arms Have Better Employment Outcomes Than Those in Usual Services

The fidelity measures of the first 2 years of service delivery indicated that IPS SE fidelity across all 30 sites was “good” on average. In-depth interviews with participants, providers, and reporting from the implementation team further bolster the findings of IPS SE fidelity measures. While it was not possible to assess fidelity formally for the final year of service delivery because the scale relies on measuring in-person service delivery (COVID-19 precluded in-person interviews), SED sites worked closely with the implementation team to continue to deliver high quality services remotely to SED participants. As such, employment outcomes were better for the treatment arms than for the Usual Services (control) arm of the SED.

It is also notable that fidelity improved over time. At the 6-month mark (YR1) average fidelity ratings were “fair,” but by the following year (YR2), ratings improved to “good.” The increase in fidelity is likely due to the technical assistance efforts of the SED implementation team and skill development among SED team members. Hence, employment outcomes in both treatment conditions also improved over time. Furthermore, the improvement of employment and earnings over time may also be related to the fact that participants began services in crisis, and many felt they needed help before they could begin to work toward employment goals. Once the sites had time to work with participants and stabilize their conditions, such as receiving treatment for previously un- and under-treated physical and mental impairments, and obtaining access to safe housing, reliable transportation, and other necessities, they were more likely to focus on employment-related goals.

6.3 Neither Treatment Condition Showed Any Impact on Disability Denial Appeal Attempts or on Disability Allowance Rates Over the 3-Year Study Period

SSA wanted to know whether successful employment intervention with denied disability applicants would reduce the number of appeals and subsequent allowances on the disability rolls. The logic appears to be something to the effect that if the SED were successful in helping more treatment group members improve their health (physical and mental) or get jobs than members of the control group, those numbers would be large enough to impact the number of denied applicants who file appeals or receive a disability allowance. We assume that higher numbers of job holders and higher earnings would result in lower numbers of appeals and allowances, effectively reducing agency costs associated with administrative processing and annual payments to allowed beneficiaries.

The study results show that both treatment groups significantly outperformed the Usual Services (control) group in both employment and earnings over the 3-year study period, as well as annually, with especially larger gains in year 3. However, those differences did not result in a commensurate impact on either allowance rates or appeal attempts. Allowance rates were roughly 15 percent in each of the three study groups, and appeals were made by nearly half of all study participants in each of the three study groups.

These results suggest that holding a job and receiving disability benefits may not be mutually exclusive events. Our qualitative data support this view. A number of study participants that we interviewed reported that their preference was to get on benefits and then also work. Further, appeals are time-limited by SSA. Denied participants have only 60 days to appeal their denial. Many participants joined the study, filed appeals, and did not seek services from the treatment teams. Only later in their tenure did they return to seek treatment. While we do not know for sure, we can speculate that many of them sought to improve their health before seeking jobs, or they waited until they received a new decision before moving forward with employment services. Several participants that we interviewed received an allowance and then sought employment to supplement their benefits or find social interaction.

The logic that having a disability and working through it to find meaningful work would reduce the need for disability benefits generally seems intuitive. However, it appears to be much more complicated than that. Given the unexpected circumstances in their lives based on the complex challenges we found this population experienced, participants most often voiced a dual goal of finding employment and at the same time seeking benefits through SSA disability programs. Several interviewees suggested that this would improve their chances to reach a stable working life. In addition, while sites provided vocational and mental health supports to participants, they did not, at any point, actively discourage them from applying for disability benefits. Such behavior was never included in any goals or directives provided to site teams. In fact, disability benefits counseling was available to all participants through both treatment programs. From this perspective, it may not be surprising that the interventions had no immediate effect on disability awards.

6.4 Outcomes for Basic-Services and Full-Services Participants Did Not Differ Significantly

The central feature of the intervention model for the Full-Service and Basic-Service treatment groups is the integration of SE with behavioral health treatment, following the evidence-based IPS model of employment services. The only difference between the two treatment arms was the availability of services from an NCC for participants in the Full-Service treatment group. While SED team leads and other service providers most often spoke in glowing terms about the successes achieved by NCCs in their work with participants, NCC efforts on behalf of a few participants per site may not be quantitatively significant. NCCs frequently seemed to spend the most time and effort with those Full-Service participants whose physical and mental illnesses were among the most complex, serious, and under-treated. The beneficiaries of NCCs' intensive efforts were usually only a few, or at most, a handful, of participants at the sites. This is not because NCCs chose to work with some participants over others but rather because these individuals were among those Full-Service participants who needed, and were interested in receiving, this sort of intensive help.

Furthermore, not all NCCs were able to dedicate time to this sort of intensive work. Time dedicated to NCC work varied widely across sites, with some teams' NCC dedicated to as little as 0.15 to 0.4 FTE. In addition to differences across sites in the quantity of NCC and medication-related services available to Full-Service participants, the quality of NCC services delivered may have varied widely across sites. Challenges to consistency across sites included insufficient hours allocated to SED work by site administrators but also turnover among NCCs, lack of clarity about the role for NCCs hired to replace the original employee, and difficulties providing adequate supervision to NCCs.

In addition, the NCC role providing medication evaluation and support was less central to the Full-Service team's service delivery than anticipated by the study design. Many participants did not use psychiatric medications and did not want or need them. Among participants who did take psychiatric medications, some received them from a prescriber not affiliated with the SED site, with whom NCCs had mixed success implementing all aspects of medication evaluation and support. Evaluation by the NCC was more consistently delivered than was treatment support, which was not requested or even indicated in many instances.

Plausibly, some team leads believed that the NCC's most important successes were in helping Full-Service participants improve their health-related behaviors through dogged persistence. NCCs helped Full-Service participants access needed health services; repeatedly reminded and encouraged participants to attend appointments; advocated for participants with their primary care, specialty, and mental health providers; discussed and explained physician recommendations to participants; and repeatedly followed up with participants to remind and encourage them to adhere to treatment.

6.5 Secondary Outcomes Associated with Clinical Recovery, Mental Health Status, and Quality of Life Were Similarly Impacted Across All Study Groups. Only Physical Health Status Showed Signs of Differential Impact of the Treatments

Average scores on clinical recovery, mental health, and quality of life all improved substantially with each study year for individuals. However, those improvements were not confined to any particular study group. Symptoms (as measured by the CSI) decreased in each of the 3 study years; mental health status (as measured by the SF-12) improved in each of the 3 study years; and quality of life (as measured by the Quality of Life scale) improved on average over each year of the study. There were no significant differences between study groups on any of these outcomes; they all improved in the same direction and with nearly identical magnitudes. The only notable difference on any secondary outcomes between any of the groups was found with physical health status (as measured by the SF-12) in the Year 2 measurements. In that case, both treatment groups showed better physical health status than that reported by the Usual Services control group in Year 2.

These results are somewhat surprising as previous studies of IPS showed significant improvement in mental health and quality of life along with employment gains, and no differential improvement in physical health status (Frey et al., 2011). The contradiction is most likely due to differences between the study populations of previous IPS studies and this unique population of denied disability applicants.

6.6 The Effect of the COVID-19 Pandemic on Outcomes Is Ambiguous

It is difficult to attribute outcomes results to the pandemic with any strong inferences since the SED was not well-positioned to draw conclusions. During the final 2 years of the demonstration, all 30 treatment providers shifted to remote—rather than in-person—delivery because pandemic-related mandates prevented in-person appointments from taking place. All previous studies setting the evidence base for IPS SE have involved in-person delivery of services. While the implementation team worked closely with sites to deliver the best remote services possible under pandemic conditions, the impact of shifting services to a remote mode of delivery on the effectiveness of IPS is unknown. It is possible that SED outcomes related to participant employment may be less positive than they would be had the sites been able to continue to deliver services face-to-face.

SED staff members described remote IPS SE service delivery as introducing new challenges to participant engagement for the purposes of accessing and using remote technology; maintaining privacy; and developing rapport. These additional challenges may account for some of the drop-off in service usage[7] seen across the 36 months of participation, which may ultimately affect outcomes negatively. However, SED providers reported that participants did not uniformly disengage due to the pandemic and related mitigation mandates. Instead, the pandemic seemed to influence engagement positively for some participants who re-engaged with their providers after a period of

---

[7] During the second month of enrollment, 55% of treatment-arm participants attended an appointment with an IPS specialists. After 1 year, 30% of participants met with their IPS specialists, and by the 2-year mark, only 15% did.

inactivity, or who became more committed to working toward their goals. Therefore, any assertions that pandemic-related service delivery challenges caused a drop in service usage are perhaps spurious.

SED service providers identified other challenges that may have affected the implementation of IPS SE services caused by the pandemic. As above, the impact of the pandemic challenges on outcomes is unclear.

- It is clear that employment and earnings among the Usual Services participants dropped precipitously compared to the treatment groups during the pandemic. With Usual Services control group employment dropping and treatment group employment remaining stable, the significance between the groups increases. We assume that the Usual Services participants faced new barriers with the pandemic that were offset by IPS services in the treatment groups.

- A temporary decrease in participant employment (and available jobs) at the beginning of the pandemic across sites and a shift in the types of employment available. In general, providers were unsure whether the pandemic would have any major effect on the overall number of participants who obtained employment. Some participants who had employment at non-essential businesses (e.g., retail, food service, hospitality) lost jobs. Participants whose unemployment benefits under Coronavirus Aid, Relief, and Economic Security Act (CARES Act) were higher than their previous wages were reluctant to search for employment that would provide an income less than their unemployment benefits. Participants feared infection, and possible death or long-term debility, should they take employment interfacing with the public.

- Local job markets altered in response to the pandemic. The pandemic may have slowed participant progress to employment, but service providers felt that they were eventually able to find suitable jobs for participants who wanted them. While non-essential businesses had reduced employment opportunities, demand for workers grew in other businesses and industries, including at grocery stores and warehouses, and in construction, healthcare, and delivery services. More full-time work in these industries became available because many people were hesitant to take jobs that put them into contact with the public, increasing their exposure to infection. The pandemic also created new jobs, such as contact tracer and "screeners" who measured the temperature of people entering buildings.

- The number and quality of contacts of IPS specialists with hiring managers suffered with remote job development. Hiring managers were less likely to respond to overtures from IPS specialists by phone than in person. Even when IPS specialists received a response to their cold call, they found that building relationships with hiring managers remotely was more difficult.

- Changes in local job markets meant that participants were more likely to receive assistance with obtaining a low-quality job than employment in a chosen career in 2020. SED staff at almost a quarter of sites (n = 7; 23%) said they shifted focus from helping participants build careers that interested them toward helping participants meet immediate needs for cash through "survival jobs"—that is, poorly paid work with no benefits and little potential for advancement. Only 40 percent of sites said that they were able to maintain a focus on placing participants in work tailored to participants' skills and interests.

7. Summary of Findings, Limitations, and Policy Implications

The SED was a multicomponent intervention aimed at improving the employment outcomes of Social Security disability applicants alleging a mental impairment who were recently denied benefits. The primary question that SSA sought to answer was whether offering the IPS model of employment services along with behavioral health and other services would foster employment and clinical recovery that leads to self-sufficiency, improved quality of life, and less demand for disability benefits. Findings from this study can help policymakers improve existing programs that may lead to increased participation of individuals with disabilities in the workforce.

This chapter revisits the original research questions posed by SSA, identifies study limitations, presents a summary of the key findings, and consolidates the findings and information from the previous chapters into a set of policy implications. The policy implications section intends to give a deeper understanding of what the study has to say about return to work for the population with an initial denial on a SSA disability benefit application.

7.1 Key Impact Findings

SSA's disability programs provide benefits to individuals who have developed a medical condition that prevents substantial work activity based on certain qualification requirements. Over the years, SSA designed various demonstrations to reduce dependence on disability benefits among beneficiaries. However, once they enter the Disability Insurance (DI) program, few beneficiaries find employment and exit the rolls (Socials Security Office, 2017). The motivation for the SED comes from the hypothesis that interventions targeted to disabled individuals before they apply for SSA benefits may achieve greater success. SED provided SSA with the first opportunity to test an early intervention through evidence-based employment services, along with behavioral health and other supports oriented toward socioeconomic determinants of health, in supporting disability benefit applicants with alleged psychiatric impairments to return to work and reduce dependence on disability benefits. In 2016, SSA released a request for proposals to conduct the SED. The RFP described the following questions related to the impact of the study:

1. To what extent does delivering appropriate, coordinated mental health treatment and employment supports lead to better employment outcomes among people with mental illness?
2. What was the impact of the demonstration for the treatment group as compared to the control group on:
    A. The proportion and timing of SSA disability applications, appeals, re-applications, and awards,
    B. SSA disability payments,
    C. Employment, hours of work, wages, and earnings,

   D. Health, access to healthcare, and healthcare costs, and
   E. Well-being and quality of life.

3. What specific programmatic changes can be made to support the efforts of people with mental illness in their effort to sustain competitive employment?

Findings presented below represent analyses conducted on data collected from participant surveys, demonstration sites, and extant sources during the study.

7.1.1 Employment and Earning Outcomes

The following findings represent analyses of interview data on the employment rate and earning outcomes:

1. Full-Service and Basic-Service participants had significantly higher rates of employment (74.0% and 74.4%, respectively) than the Usual Services participants (64.1%) during the 3-year study period.

2. The annual employment rate for the Full-Services group was lowest in Year 1 (53.9%), but increased in Year 2 (60%) and remained above Year 1 in Year 3 (54.3%). The Basic-Services and Usual Services groups were similar in their pattern of performance across study years with both groups achieving their highest employment rates in Year 1 (58.1% and 50.1%, respectively) and their lowest employment rates in year 3 (53.2% and 42.9%, respectively).

3. Full-Service and Basic-Service participants had significantly higher weeks of employment (45.9 and 46.5 weeks, respectively) than the Usual Services participants (37.0 weeks) during the 3-year study period.

4. Treatment group participants reported significantly higher number of hours worked during the study period. Full-Service and Basic-Service participants had 1,289 and 1,308 hours worked, respectively, and Usual Services participants had 1,045 hours worked during the 3-year study period.

5. There is no significant difference between treatment arms and control in terms of the percentage of workers earning SGA in the past month. This ratio increases each year in all three study arms reaching to 16.5, 15.5, and 13.9 percent for Full-Services, Basic-Service, and Usual Service categories, respectively.

6. Over the entire 3-year study period, more than one-third of workers were designated steady workers. Both treatment groups had significantly higher percentages than the control group (Full-Services, 37.26%; Basic-Services, 37.17%; Usual Services, 30.97%).

7. Treatment group participants reported significantly higher earnings during the study period. Full-Service and Basic-Service participants earned $17,925 and $17,556, respectively, and the Usual Services participants had earnings of $13,547 during the 3-year study period.

8. Further analyses of earnings indicate that treatment participants who enrolled later in the enrollment period (from June 2018 through March 2019), on average, saw larger earnings gains compared to the Usual Services control group participants who enrolled in the same

quartiles compared to the early in the enrollment period (from December 2017 to June 2018).

7.1.2 SSA Benefits

1. Among those enrolled in the SED, approximately 15 percent (439 participants) received an allowance for disability benefits during the 3 years of enrollment in the study. There were no significant differences in allowance rates between the study arms.

2. There were no significant differences in the decision to appeal or the number of appeals made by participants in each study arm. Almost half (45-46%) of participants filed at least one appeal during the study period. Among those who filed an appeal, more than half (55-60%) filed only one appeal. The average number of appeals filed among those who appealed at all was 1.7.

3. Among those approved for disability, the average total disability payments made during the study was approximately $30,000 per approved participant. The treatment did not have a significant impact on these average amounts.

4. The characteristics that significantly predict allowance rates are age, health status, and work history. Older study participants, participants with lower SF-12 PCS (at baseline), and people who were not working at baseline were more likely to receive an allowance.

5. Local area economic and health data also had a significant relationship with allowance rates. Those living in states with higher levels of hospitalizations due to COVID-19, for example, were more likely to receive an acceptance to the disability rolls.

7.1.3 Health and Quality of Life

1. On average, study participants showed significant improvement in mental health each year of the study. However, when comparing the study arms, the improvements in mental health status were not significantly larger for Full-Service or Basic-Service participants compared to Usual Services.

2. By the end of the third year in the study, participants in each study arm saw an average improvement of more than 5 points in SF-12 MCS, representing approximately a one-half standard deviation increase.

3. Each study arm showed improvements in SF-12 PCS compared to baseline. Comparing the study arms, the improvements in physical health status were not significantly larger for Full-Service or Basic-Service participants compared to Usual Services by the end of the third year in the study.

4. The magnitude of the differences is between 1 and 2 scale points, on average. In the third year of study enrollment, we see larger gains in PCS for the Full-Service relative to the Basic-Service participants.

5. When comparing the study arms, we see no significant differences in the changes in quality-of-life scores at each annual survey over the course of the study.

7.1.4 Utilization of Healthcare Services

1. Participants in the Full-Service and Basic-Service groups were each more likely than the Usual Services participants to complete preventive care visits during their enrollment in the study.

2. When comparing the study arms, we see no significant differences in the total number of ER visits and hospitalization over the course of the study.

3. Participants in the Full-Service group had slightly fewer hospital stays and ER visits for mental problems compared to the Basic-Service group.

These findings clearly answer SSA's first question showing that both interventions lead to better employment outcomes among of individuals who apply for disability benefits for a mental health impairment and receive an initial denial of benefits. Both treatment groups attained significantly better outcomes compared to the control group in terms of employment and earnings outcomes. Full-Services and Basic-Services groups also experienced improvements in other outcomes, including mental health, physical health, and quality of life, but the changes were not significantly different when compared to Usual Services group.

7.1.5 Cost-Benefit Findings

The analyses indicate that the benefits did not outweigh costs during the intervention period for either the Full-Service or Basic-Services participants. However, we believe the net benefit figure understates the net benefits of the intervention when a slightly longer timeframe is employed. There is evidence from follow-up studies of IPS interventions that benefits from work extend beyond the initial period of intervention (Becker et al., 2007; McHugo et al., 1998). What is more, trends in employment and service utilization provide further suggestive evidence that the SED intervention may become cost effective in the years beyond the initial 3-year study period. SED treatment group participants saw greater gains in employment and earnings relative to the Usual Services group in Year 3 of study enrollment compared to Years 1 and 2. At the same time, utilization of IPS and care management services among treatment group participants dropped on average in the third year of enrollment compared to the first year. Combined, these trends suggest that employment benefits may continue as engagement with services drop-off. However, whether these benefits continue for individuals who no longer receive SED services is an open question that merits further investigation. The difference in employment outcomes between earlier and later enrollees also suggests that fully established programs may generate greater economic benefits than those in the startup phase and still learning to serve denied disability applicants. Therefore, established programs that have experience serving this population may be more cost effective.

7.2 Policy Implications

The findings from the study have implications for policies related to SSA's disability benefit programs, as well as for employment policies of other federal agencies as these pertain to individuals with a mental health impairment. Providing access to all the services tested in the SED likely would be beyond SSA's authority. The services implemented in the study should be of interest to many federal agencies that serve individuals with mental health impairments. The purpose of this section is to provide a discussion of the potential for the SED findings to inform policy discussions at SSA and other federal agencies. The intent of the discussion is to provide answers to the remaining SSA research questions about programmatic disincentives and programmatic changes needed to support return-to-work efforts for applicants with an initial rejection for disability benefits.

What programmatic disincentives create barriers for applicants with alleged mental illness to return to work?

Working Under the Level of SGA. The study findings presented above make it clear that a substantial percentage of applicants would consider participating in an SED-like program. In 2019, SSA denied SSDI and/or SSI benefits to over 2 million applicants. Among those adjudicated for a diagnosis, approximately 12 percent have a mental disorder. As a rough estimate, the number of denied applicants with a mental disorder in a given year who may be eligible for an intervention such as the SED is approximately 250,000 (Social Security Office, 2020a; Social Security Office, 2020b). With an enrollment rate of 25 percent, we may roughly estimate 72,500 applicants taking up the offer to participate in an SED type of intervention. Taking the finding further, the data suggest that both interventions resulted in about a 10 percent increase in employment rate relative to the control group, which leads roughly to 7,250 more participants finding employment due to the mental health and vocational services support.

However, only about one-third of individuals who obtained employment earned over the limit of SGA. Qualitative interviews indicated that a sizable number of participants are continuing their disability benefit applications and they are hesitant to work in jobs that pay them over the SGA because it may negatively impact their chances of receiving disability benefits. The typical disability benefit applicant in this study is definitely interested in obtaining and maintaining employment but would also like to receive certain level of disability benefits until they reach self-sufficiency. Returning to work with no reliable support through disability benefits can make them vulnerable to economic shock in their life such as missing utility bills, rent payments, and other expenses. These temporary changes in their income may have long-lasting consequences, including homelessness, child welfare involvement, or hospitalizations.

No Referral to Supports/Services After Initial Denial for Disability Benefits. When an applicant receives a denial on the disability benefit application, there is currently no system in place to provide referrals to health supports and vocational services. The experience of the SED suggests that when offered SED-type support services, on average, 25 percent of applicants would be willing to sign up for the programs. Community agencies providing mental health and employment support services would be ideal places to refer applicants who receive initial denial on their applications. This study makes it clear that significantly more individuals can find and maintain employment that results in higher earnings if they participate in SED-type programs.

What are the other barriers for applicants with mental illness to return to work?

The Key Role of Assistance in Healthcare Management. The SED participants had a complex set of needs for physical and mental healthcare services. The NCC role was a valuable addition to the Full-Service teams in helping some participants improve health behaviors, especially behaviors related to compliance with treatment for chronic impairments. There is an obvious need for healthcare management support to help participants navigate through the complex healthcare system, make and keep their doctor appointments, and communicate their needs to providers effectively. Participants did not utilize intensive services such as active medication management of their psychiatric medications through the NCC as often as we expected. This was in large part due to many participants not using psychiatric medications and asserting they did not want or need them. The majority of participants relying on SED funding for healthcare came from the catchment areas in non-Medicaid expansion states.

Need for Care Management. The findings indicate that the participants have complex challenges requiring wraparound services and supports in addition to the employment services they received from the sites. Providing ongoing care management was crucial for remedying participants' immediate needs. Care managers provided referrals and coordination across all basic areas of need. Most often, care managers helped participants cope with life crises by providing them with resources sufficient to attenuate their hardships so that they could focus on long-term solutions to their health and economic challenges. These services were also important in keeping focus on individualized treatment plans and engagement of participants in intervention services. Slightly less than half of all participants received assistance with housing; more than half received help with medical care; and more than 60 percent received help managing symptoms of physical and/or mental illnesses.

Need for Financial Support During Critical Times. The Westat team processed a total of 10,021 reimbursement claims totaling to $2,846,107 during the 54 months between December 2017 and June 2022. These disbursements directly supported study participants working toward employment or to overcome challenges to finding and maintaining employment. Sites reported that providing financial support for certain unexpected needs such as temporary rental assistance, transportation, auto repair, medication co-pays, dental bills, glasses, and similar needs were critical in removing barriers to employment for participants when no other resource in the community would have paid for these items. SSA currently reimburses beneficiaries for impairment-related work expenses (IWRE), but there is no mechanism in place for people who are at risk to get on disability rolls.

7.3 Study Limitations

This section discusses external validity and the extent to which the SED findings generalize to the target population of individuals who apply for disability benefits for a mental health impairment and receive an initial denial of benefits, and to the existing system of community mental health agencies in the United States. This section also discusses the internal validity of the study and the extent to which the findings can support inferences about the effectiveness of the intervention, and the likelihood that changes in policy or programmatic guidance would produce similar results.

7.3.1 External Validity

There are two main topics of concern about external validity. One topic is selection of the study sites and their representativeness of existing community mental health centers in the United States. The other is the selection of disability applicants with initial denials and their representativeness of the population of this group in the United States. This latter question concerns the extent to which beneficiaries who participated in the SED represent the universe of disability applicants with alleged mental health disabilities who received initial denials.

7.3.2 Site Representativeness

Thirty community agencies in cities across all seven SSA regions and in 20 U.S. states comprised the setting for the demonstration. The interventions tested in this demonstration included evidence-based IPS SE services. Thus, the community agencies recruited for the study were already familiar with IPS and they were all capable of having the capacity to take part in these complex interventions. Further, they were willing to establish new services to meet the needs of the SED treatment intervention. As a result, the study sites were not typical of community service organizations across the country, but neither were they atypical, such as the academic community health centers offering highly specialized services.

The study did not use a random selection of community agencies across all regions of the United States. Such a selection process might have increased the external validity of the study, but based on considerable implementation evidence from earlier demonstrations, the study team concluded that it was not feasible given the requirements of the demonstration services. In addition, because not all sites would have the capability to provide access to the required treatment intervention services, there was a possibility that random selection of sites would result in some sites failing completely. There were several other factors that affected site selection decisions, including (1) the expected number of denied disability applicants that a typical community agency could serve for the study period; (2) the potential number of denied disability applicants in specific agency catchment areas that we could recruit for the study; and (3) the generalizability of the study results to broader geographic regions. In addition, we wanted sites to represent both Medicaid expansion and non-expansion states. Selected sites reflect a national diversity including a balance of urban and rural sites, as well as sites that reflected ethnic and racial diversity.

7.3.3 Beneficiary Representativeness

The SED enrolled 3,000 recently denied disability applicants living within the catchment area of one of the 30 community agencies contracted by the study to deliver intervention services. It is important to understand who was interested in participating in the SED, who eventually agreed to join the study, and who remained engaged throughout. Several selection points in the recruitment process introduced potential sources of bias and threats to external validity. The investigators made an effort to understand these potential sources of bias and correct. We compared basic demographics and other relevant variables between eligible candidates selected for recruitment and those eligible candidates not selected for recruitment. for them, where possible.

The two groups were very similar in demographic characteristics (age, gender, and education), as well as job history, number of jobs held in the past 15 years, and weekly pay of most recent job with no significant differences noted. The two groups differed on two variables related to their disability status, but the overall numbers were relatively small. Random assignment of the original 3,000 study participants resulted in 994 study participants assigned to the Full-Service treatment group; 1,004 assigned to the Basic-Service treatment group; and 1,002 assigned to the Usual Services (control) group. The resulting assignments revealed no differences between the study arms on key variables, including age, age grouping, gender, education, work history, health insurance status, marital status, arrest history, physical or mental health scores on the SF-12, or household income in the past month. One variable—race and ethnicity—did appear as significant across the three study arms. Fewer Whites participants received assignments to the Basic-Service group, while more Black and more Hispanic participants received assignments to that group compared to either the Full-Service treatment group or the Usual Services (control) group.

7.3.4 Internal Validity

The SED employed random assignment of beneficiaries within sites to the treatment or control group. This procedure alone increased the internal validity of the observations and the ability to make inferences about the effectiveness of the treatment intervention. One key characteristic of the study design was that the treatment intervention was the intent-to-treat (ITT) design feature. The ITT focuses the interpretation of the results on the entire sample of study participants—including all beneficiaries who had access to the treatment intervention services without regard to the extent to which they might have participated in those specific services. The overall findings did not differ significantly with the results presented in this report and with the ITT deign findings.

One potential issue is that we selected participants from the SSA list of individuals who have recently received an initial denial on their disability benefit application. In most cases, these individuals were not receiving mental health services and vocational rehabilitation from these community agencies. This may have introduced other potential influences on the results. Recruiting individuals with less connection to the community mental health service system might mean worse outcomes than expected from prior studies of IPS SE.

References

Becker, D., Whitley, R., Bailey, E., and Drake, R. (2007, July). Long-term employment trajectories among participants with severe mental illness in Supported Employment. *Psychiatric Services*, 58(7): 922-28.

Bond, G.R., Drake, R.E., and Becker, D.R. (2012). Generalizability of the Individual Placement and Support (IPS) model of supported employment outside the US. *World Psychiatry*, March 2013, 11(1): 32-39.

Brinchmann, B., Widding-Havneraas, T., Modini, M., Rinaldi, M., Moe, C.F., McDaid, D., Park. A.-L., Killackey, E., Harvey, S.B., and Mykletun, A. (2019, November). A meta-regression of the impact of policy on the efficacy of Individual Placement and Support. *Acta Psychiatrica Scandinavica*, 141(3): 206-220.

Freeman, W.J., Weiss, A.J., and Heslin, K.C. (2018) *Overview of U.S. hospital stays in 2016: Variation by geographic region* (Statistical Brief #246). Rockville, MD: Agency for Healthcare, Research and Quality. Available from: https://www.hcup-us.ahrq.gov/reports/statbriefs/sb246-Geographic-Variation-Hospital-Stays.jsp.

Frederick, D.E., and VanderWeele, T.J. (2019, February). Supported employment: Meta-analysis and review of randomized controlled trials of individual placement and support. *PLOS ONE*, 14(2).

Frey, W., Drake, R., Bond, G., Miller, A., Goldman, H., Salkever, D., and Holsenbeck, S. (2011). *Mental health treatment study final report* (Report prepared for the Social Security Administration). Available from: https://www.ssa.gov/disabilityresearch/documents/MHTS_Final_Report_508.pdf.

Jans, L., Stoddard, S., and Kraus, L. (2004). *Chartbook on mental health and disability in the United States* (An InfoUse Report). Washington, DC: U.S. Department of Education, National Institute on Disability and Rehabilitation Research. Retrieved October 21, 2009. Available from https://silo.tips/download/chartbook-on-mental-health-and-disability-in-the-united-states.

Kinoshita, Y., Furukawa, T.A., Omori, I.M., Marshall, M., Bond, G.R., Huxley, P., and Kingdon, D. (2013). Supported employment for adults with severe mental illness. *Cochrane Database of Systematic Reviews,* 2010(9), CD008297. DOI: 008210.001002/14651858.CD14008297.

Kjellsson, G., Clarke, P., and Gerdtham, U. (2014, May). Forgetting to remember or remembering to forget: A study of the recall period length in health care survey question. *Journal of Health Economics*, 35: 34-46.

Marshall, T., Goldberg, R.W., Braude, L., Dougherty, R.H., Daniels, A.S., Ghose, S.S., George, Preethy, and Delphin-Rittmon, M.E. (2014). Supported employment: Assessing the evidence. *Psychiatric Services*, January 2014, 65(1): 16-23.

McHugo, G.J., Drake, R.E., and Becker, D.R. (1998). The durability of supported employment effects. *Psychiatric Rehabilitation Journal*, 22(1), 55-61.

Metcalfe, J.D., Drake, R.E., and Bond, G.R. (2018, January). Economic, labor, and regulatory moderators of the effect of Individual Placement and Support among people with severe mental illness: A systematic review and meta-analysis. *Schizophrenia Bulletin*, 44(1): 22-31.

Modini, M., Tan, L., Brinchmann, B., Wang, M., Killackey, E., Glozier, N., Mykletun, A., and Harvey, S.B. (2016, July). Supported employment for people with severe mental illness: Systematic review and meta-analysis of the international evidence. *British Journal of Psychiatry*, 209(1): 14-22.

Moore, B., and Liang, L. (2020). *Costs of emergency department visits in the United States, 2017* (Statistical Brief #268). Rockville, MD: Agency for Healthcare, Research and Quality. Available from: https://www.hcup-us.ahrq.gov/reports/statbriefs/sb268-ED-Costs-2017.jsp.

Riley, J., Drake, R.E., Frey, W., Goldman, H.H., Becker, D.R., Salkever, D., Marrow, J., Borger, C., Taylor, J., Bond, G.R., and Karakus, M. (2021, May). Helping people denied disability benefits for a mental health impairment: The supported employment demonstration. *Psychiatric Services*, 72(12): 1434-40.

Social Security Office of Retirement and Disability Policy (2017). *Annual statistical report on the Social Security Disability Insurance Program, 2017.* Available from: https://www.ssa.gov/policy/docs/statcomps/di_asr/2017/di_asr17.pdf.

Social Security Office of Retirement and Disability Policy (2019). *Annual statistical report on the Social Security Disability Insurance Program, 2019.* Available from: https://www.ssa.gov/policy/docs/statcomps/di_asr/2019/index.html.

Social Security Office of Retirement and Disability Policy (2020a). *Annual statistical report on the Social Security Disability Insurance Program, 2020.* Available from: https://www.ssa.gov/policy/docs/statcomps/di_asr/2020/sect04.html.

Social Security Office of Retirement and Disability Policy (2020b). *SSI annual statistical report,* 2020. Available from: https://www.ssa.gov/policy/docs/statcomps/ssi_asr/2020/sect10.html.

Suijkerbuijk, Y.B., Schaafsma, F.G., van Mechelen, J.C., Ojajärvi, A., Corbière, M., and Anema, J.R. (2017, September). Interventions for obtaining and maintaining employment in adults with severe mental illness, a network meta-analysis. *Cochrane Database of Systematic Reviews,* 9. Available from: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD011867.pub2/full.

Taylor, J.A., Salkever, D., Frey, W., Riley, J., and Marrow, J. (2020). *Supported employment demonstration: Final enrollment analysis report.* Rockville, MD: Westat. Available from: https://www.ssa.gov/disabilityresearch/documents/SED-FinalEnrollmentAnalysisReport.pdf.

Taylor, J.A., Salkever, D.S., Frey, W.D., Riley, J., and Marrow, J. (2021, November). Enrollment in the supported employment demonstration: An employment intervention for denied disability benefits applicants with a mental impairment. *Administration and Policy in Mental Health Services Research*, 49(6): 909-26.

U.S. Bureau of Labor Statistics (2020a). *BLS employer costs for employee compensation: Private industry dataset.* Available from: https://www.bls.gov/bls/news-release/ecec.htm.

U.S. Bureau of Labor Statistics (2020b). *BLS National Compensation Survey (NCS).* Available from: https://www.bls.gov/eci/.