Behavioral Interventions to Improve Work Search Among UI Claimants: Results from North Carolina and Washington

February 2022

Samia Amin, Greg Chojnacki, Bill Congdon, Scott Davis, Andrew Langan, Jonah Deutsch, Erin Welch, Ariella Spitzer, Amy Johnson

This page has been left blank for double-sided copying.

Acknowledgments

This page has been left blank for double-sided copying.

Contents

Exhibits

I.   Introduction

A.   Importance of improving work search and reducing UI improper payments

Timely work search is critical for helping individuals return to work. Actively looking for work soon after losing a job can be important because the likelihood of finding a job decreases the longer an individual is unemployed (Faberman and Kudlyak 2019). To encourage a rapid return to work, the Unemployment Insurance (UI) program requires that individuals receiving UI benefits actively look for work while receiving benefits.[1]

Improving work search among UI claimants has been an important strategic priority for the U.S. Department of Labor (DOL). DOL prioritizes work search both to achieve its mission of helping workers and improve compliance with federal legislation designed to improve the integrity of public programs (Exhibit I.1). DOL's Chief Evaluation Office (CEO) sought to understand how behavioral interventions might be used to improve adherence to work search requirements among UI claimants. Behavioral interventions apply findings from behavioral science about how people make and act on decisions to improve the design of public policies and programs. Since 2014, CEO has partnered with many different DOL agencies through the Department of Labor Behavioral Interventions (DOLBI) project and similar efforts to improve labor outcomes.[2] In 2019, the DOLBI project focused its efforts on improving work search compliance in partnership with the Office of Unemployment Insurance (OUI) and two interested states that were motivated to experiment and learn about how behavioral interventions might be used to reduce improper payment rates—North Carolina and Washington.

Exhibit I.1. DOL's strategic prioritization of reducing work search errors and improper payments

The Improper Payments Elimination and Recovery Act (IPERA) of 2010 requires federal agencies, including the Department of Labor (DOL), to limit improper payments to less than 10 percent of total payments. IPERA defines improper payments as "any payment that should not have been made or that was made in an incorrect amount (including overpayments and underpayments)." When the present study was launched in 2018, the UI improper payment rate was 12.95 percent. Work search errors were the leading contributor to UI overpayment rates, accounting for 40.51 percent of dollars overpaid across all states in 2018. Reducing improper payments resulting from work search errors as quickly as possible is an important strategic priority for DOL. Recent investments to support this priority include DOL-funded provision of intensive technical assistance services to states to help reduce improper payments, with a special focus on states DOL designated as high rate/high impact.

Note: Improper payment rates are measured by performance year and consist of both over- and underpayments. The 2018 performance year covered July 2017 to June 2018. The reported improper payment rate is from the Benefit Accuracy Measurement State Data Summary for Improper Payment Information Act Performance Year 2018, available at https://oui.doleta.gov/unemploy/bam/2018/IPIA_2018_Benefit_Accuracy_Measurement_Annual_Report.pdf.

---

[1] It's a statutory requirement that all individuals look for work, though there are other statutory exceptions to this, such as approved training and Short-Time Compensation.

[2] The U.S. Department of Labor (DOL) Behavioral Interventions project explores how insights from behavioral science can be used to influence how people make and act on decisions to improve the performance and outcomes of DOL programs. It is sponsored by the DOL Chief Evaluation Office and conducted by Mathematica, American Institutes for Research, ideas42, the Urban Institute, and the W.E. Upjohn Institute. For more information, visit https://www.dol.gov/agencies/oasp/evaluation/topic-areas/behavioral-interventions.

B. Overview of goals and design of work search behavioral interventions and evaluations

Efforts to use behavioral science to understand work search among UI claimants can draw on a rich and growing body of research. As discussed in detail in the following chapter, three strands in existing research ground potential applications of behavioral science:

1. Literature on the behavioral dimensions of work search

2. Research on the behavioral dimensions of program compliance in general

3. Research that incorporates a behavioral perspective into economic models of UI in particular

Collectively, this body of literature suggests that many different behavioral approaches might yield benefits, and rich possibilities exist for learning what approaches work for whom, and where.

Our goal under the DOLBI project was to initiate quick behavioral interventions (BI) to address work search errors stemming from a subset of the options implied by the literature. Specifically, the study team focused on the potential for developing interventions modifying how claimants receive information about work search to address behavioral barriers (such as complexity) that reduce work search compliance. The accompanying evaluations were designed to answer the following questions:

- Can light-touch behavioral interventions—that is, behavioral interventions that can be delivered with limited resources and minimal burden on staff—increase UI claimants' knowledge of work search requirements?

- Can behavioral interventions help reduce improper payments due to work search errors?

- What can we learn about implementing behavioral interventions in the UI system?

Our experience conducting behavioral trials suggests that the first phase of experimentation yields important lessons on improving design and implementation procedures that can be incorporated for additional testing and adaptation to the local context. This work was therefore designed to generate an initial set of lessons for the field from our partnerships with two states, which then could form the basis of subsequent experimentation and learning.

To guide our work with our partner states, we used the six-step process, shown in Exhibit I.2 and developed for the DOLBI project (Darling et al. 2017), to design and implement BIs:

Exhibit I.2. Six-step process
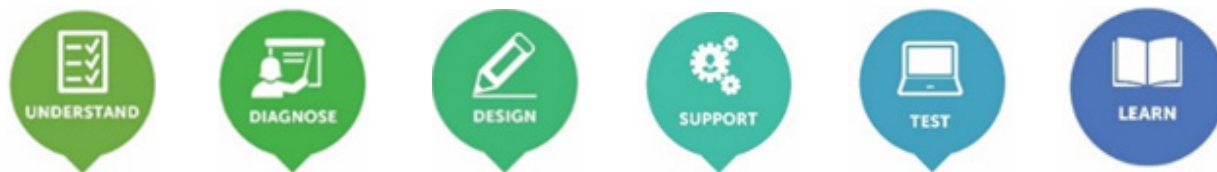
As detailed in the individual chapters, we worked closely and collaboratively with our partner states, CEO, and OUI to accomplish the following:

1. **Understand** areas in which program performance could be improved

2. **Diagnose** potential behavioral barriers (features of program design or context that lead to counterproductive decisions or behaviors among the target population)

3. **Design** interventions that address those barriers and evaluations to determine whether the interventions work

4. **Support** implementation of the behavioral intervention

5. **Test** the intervention's effectiveness

6. **Learn** from and share the findings

Exhibit I.3 summarizes the resulting behavioral interventions and evaluations in each state.

Exhibit I.3. Key features of interventions and evaluations

	North Carolina	Washington
Behavioral intervention	A **weekly pop-up alert** within the online claims filing system; hyperlink to a **tip sheet** outlining what counts as valid employer contacts and adequate proof for different modes of contact **Weekly email** with the same content	**A single email** sent after UI claimants file their first weekly claim, with links to a work search log documenting the required work search details and information on completing work search activities
Evaluation design		
Impact study	• **Stage 1**: a randomized controlled trial (RCT) focused on self-attestations of work search • **Stage 2**: a pre-post quasi-experimental design (QED) focused on work search behavior	RCT focused on increasing claimant information and simplifying record keeping
Implementation study		
Quantitative engagement data	Counts of click-throughs on links embedded in alerts and emails; acknowledgment of alerts	Email open rates; share of treatment group clicking on each link within the email
Qualitative data on implementation and study context	Phone interviews with staff and on-site focus groups with customers and staff	Phone interviews with staff

C. Summary of findings

The COVID-19 pandemic disrupted the economy, causing large spikes in unemployment and placing tremendous burden on the workload and service delivery priorities of workforce and unemployment insurance agencies across the United States. The pandemic consequently disrupted study plans in both states and limited our learning. As detailed in subsequent chapters, the onset of the pandemic truncated our analytic sample in North Carolina, limited our ability to collect outcome data in Washington State, and created conditions (such as federal legislation giving states the latitude to suspend work search requirements) that precluded additional collaboration regarding refining interventions and pursuing further experimentation and learning in both states. In Washington, additional concurrent changes to the

work search reporting process reduced the contrast between our intervention and the comparison group in terms of the degree to which claimants received salient information about work search requirements.

Despite these challenges, these initial studies have yielded the following important findings and lessons learned, which we discuss in greater detail in Chapter V. Although neither intervention appears to have increased knowledge of work search requirements or improved work search behavior prior to the onset of the pandemic, state partners reported that the interventions helped them see the value of applying behavioral science to address program challenges and created more widespread interest in behavioral approaches. In addition, this work demonstrated the feasibility of quickly implementing interventions for work search processes. It highlighted the value of state UI administration systems and data, though it also revealed some of their limitations. Finally, these partnerships helped highlight a set of best practices for collaborating with states on these kinds of engagements. Lessons related to these practices include the value of site visits to establish understanding and trust as a foundation for collaboration; the increased efficiency gained by aligning different technical assistance and research efforts; the importance of engaging with users at the diagnosis and design stages to improve intervention design; and the necessity of actively monitoring for unanticipated developments.

Insights from these studies also point to additional areas of learning, as detailed in Chapter V. They suggest that exploring changes in the intensity, mode, timing, and alignment of communications may help improve their impacts. Insights from our site visit indicate that it may be worthwhile to focus on additional behavioral barriers beyond knowledge and complexity, such as motivation and planning. Applying behavioral insights to the design and delivery of reemployment services, such as the Reemployment Services and Eligibility Assessment (RESEA) program, seems to be a promising area in which applying behavioral insights may improve program impacts. Finally, state partners' appetite for applying behavioral insights suggests potential payoffs from cultivating longer-term engagements with states. Such longer-term engagements could create opportunity to iterate on interventions and deepen learning about how to address behavioral challenges.

D.   Organization of the report

After discussing relevant aspects of the existing behavioral science literature in Chapter II, this report presents the design and findings of the North Carolina and Washington projects in Chapters III and IV, respectively. Chapter V summarizes key findings across these two projects and describes the implications for future testing of applied behavioral science to promote work search among UI claimants.

II. Behavioral Science Perspective on Work Search Among UI Claimants

Behavioral science, and economics research that incorporates a behavioral perspective, updates our understanding of how people decide and act in labor markets and interact with labor market programs (Babcock et al. 2012; Darling et al. 2019). As a result, this research also suggests the potential for intervention and policy innovation. Three strands of this research are particularly informative for the context of our two projects:

1. Research on the behavioral dimensions of program compliance in general
2. Research on the behavioral dimensions of work search in general
3. Research that incorporates a behavioral perspective into economic models of work search incentives in UI

A. Program compliance: Behavioral science insights

Expanded models of program compliance represent one of the strands of behavioral science with potentially promising applications to improve UI program outcomes related to program requirements, such as conducting accurate work search. The standard model of compliance with program requirements assumes that people rationally calculate the benefits of noncompliance, the probability of being found noncompliant, and the costs of associated penalties (Becker 1968). This model suggests—even if implicitly—that willful noncompliance is the driver of noncompliance and work search errors, and workers who fail to comply with requirements know what they are supposed to do but choose not to do it.

A behavioral science perspective emphasizes the potential role of other factors that impede compliance. Behavioral science findings suggest that, as with behavior in general, people will tend to make decisions and take actions that are compliant when rules are salient and it is easy to comply with them (Mazar and Ariely 2006; Shu et al. 2012; Hallsworth et al. 2017). This perspective has potential implications for interpreting and understanding observed responses to work search requirements in UI. For example, workers with limited attention may be unaware of requirements or their details. Even those intending to comply may experience the complexity of requirements or complexities associated with adequately demonstrating or documenting compliance as barriers.

This literature thus points to a distinct set of potential interventions or policy innovations worth investigating that might address work search errors in UI and associated improper payments. These changes could include timely, salient communications with claimants to clarify and remind them of both their search and reporting obligations. (The interventions examined in this report fall into this category.) Other innovations might address the work search requirements themselves, such as simplifying the set of allowable activities or reducing the required number of activities to be completed each week.

B. Work search: Behavioral science perspective and empirical evidence

Behavioral science brings both a perspective and evidence to help understand how and why people search for work. Standard economic search-and-matching models assume people efficiently search for openings and evaluate job offers against some minimum acceptable wage (for example, see Diamond 1982 and Mortensen and Pissarides 1994). They further assume that this wage, called the reservation wage, reflects consistent and well-informed assessments of one's own productivity and outside opportunities. These

models recognize that job search is costly, requiring time, effort, and potentially out-of-pocket costs from individuals as they go about finding new work. However, they do not account for search frictions that arise due to inconsistencies or imperfections in workers' decision making.

How people search for work

A behavioral perspective adds theoretical underpinnings—and provides both direct and indirect evidence—that people may search for work in less-than-optimal ways.

- Timing and perceptions of search effort: For example, research suggests that people may procrastinate in doing work search by putting off their efforts in ways that do not reflect their preferences to find and take work (DellaVigna and Paserman 2005). Other research finds that some job seekers underestimate the value of work search efforts, whereas others might be too optimistic about their ability to find work quickly; both situations may lead workers to underinvest in search activities and slow their return to work (Spinnewijn 2015; Mueller et al. 2021).

- Scope and focus of work search: Some studies suggest that workers may search too narrowly along a number of dimensions. For example, they might search for work in only very local labor markets (Marinescu and Rathelot 2018). Other evidence suggests workers might focus their search on a narrow set of occupations, such as those closely related to their previous job (Belot et al. 2019). Others have hypothesized that factors related to identity might affect their search. For example, male workers may neglect searching for work in occupations traditionally held by female workers—a factor sometimes called identity mismatch (Katz 2017).

- Evaluation of new job opportunities: Finally, a body of evidence suggests that reservation wage formation may be affected by factors such as reference dependence (people's tendency to evaluate choices relative to reference points); for example, workers might anchor their wage expectations to their wages in a previous job (Krueger and Mueller 2016).

For unemployed job seekers, these basic behavioral factors that affect work search are potentially heightened by accompanying psychosocial stress. The negative effects of unemployment on psychological well-being are well documented (for example, Knabe and Ratzel 2011). The emerging psychology of scarcity suggests that such effects, in combination with the negative effects of unemployment on financial security, can potentially exacerbate general behavioral tendencies such as procrastination or inattention in decision making (Mani et al. 2013; Mullainathan and Shafir 2013).

Why people search for work

Finally, standard models of work search typically embed assumptions about *why* people search for work, in addition to *how* they search. The standard assumption in labor economics is that people search for and take work primarily because they need the compensation it provides to support consumption, and that work itself is something they would otherwise prefer to avoid (Baily 1978). A behavioral perspective considers and takes more seriously the evidence showing that work provides intrinsic rewards for many people and is a source of identity or meaning (for example, Loewenstein 1999; Karlsson et al. 2004; Akerlof and Kranton 2000). The benefits of work can manifest themselves in a range of nonmonetary outcomes for workers, including physical health, psychological well-being, happiness, and subjective well-being (for example, Winkelmann and Winkelmann 1998).

Together, this literature suggests a range of potential interventions or policy innovations worth investigating that might improve the outcomes of work search among UI claimants. The possibilities

include communications with claimants that better reflect and respond to worker motivations for returning to work, or supporting those with challenges in searching for work. Information interventions could offer an orientation in messaging that assumes workers are motivated to find work. Other innovations might include those that structure program requirements and incentives in ways that are responsive to the effects of scarcity or help people overcome procrastination, either by providing workers with greater flexibility to alleviate its effects or designing program features and communications that counteract inattention and procrastination.

C.   Work search incentives in UI: Implications of behavioral factors

Unemployment insurance pays qualifying unemployed workers a portion of their pre-unemployment wage, up to a maximum benefit amount, for a limited number of weeks, while they are unemployed and searching for work. To establish and maintain their eligibility to receive UI benefits, workers must meet both monetary and nonmonetary program eligibility requirements. These requirements include having enough recent earnings and duration of employment, having lost work through no fault of their own, being willing and available to accept suitable employment, and satisfying program requirements (such as work search). Work search requirements vary by state, but typically workers claiming UI benefits must undertake and record or report a certain number of qualifying activities each week (such as making contacts with employers). When payments of UI benefits are made to workers who do not meet (or are later found not to have met) these requirements, such payments are classified as improper payments.

The classical economic theory that is often applied to understand work search and program incentives justifies the necessity of UI work search requirements (along with other features of the program's design, such as the limited duration and level of benefits) (Baily 1978). The concern under the classical model is that UI creates what economists refer to as *moral hazard*: that by insuring workers against hardship when they lose work, the program might also dull their incentives to search for and return to work. Standard models of UI delineate these trade-offs—the benefits of supporting people who are out of work and the costs of moral hazard—and inform how program design might balance them (Baily 1978; Chetty 2008). Under this model, work search requirements serve to realign the incentives of workers and mitigate moral hazard.

A behavioral perspective brings into the model features of individual decision making that mediate work search—such as procrastination or reference dependence—and reconsiders the form and extent of that trade-off (Mullainathan et al. 2012; Chetty 2015; Marinescu and Skandalis 2019). If the typical worker wants to be employed but finds work search difficult, the optimal program response might be to provide assistance to workers to help them better realize their intentions, rather than imposing strict work requirements and associated sanctions. Modern empirical economics literature finds that UI tends to have only relatively modest effects on both employment and overall unemployment levels (for example, Card et al. 2007; Rothstein 2011; Marinescu 2017; Chodorow-Reich et al. 2019). This failure to find consistent evidence of UI creating a strong moral hazard is at least broadly consistent with a model in which behavioral factors play an important role in mediating the effects of UI on work incentives.

Taken together with the research on work search and program compliance, a behavioral perspective indicates the importance of using a broadly empirical approach when designing and implementing program elements such as work search requirements. This approach helps to improve compliance with the requirements but also ensures that these requirements serve the underlying objectives of the program. When it is possible to measure longer-term outcomes such as unemployment duration, it is important to assess both whether a program change improves compliance with requirements, and whether it helps

unemployed workers return to work more quickly. This approach suggests a number of potential interventions or policy innovations worth investigating that might improve the work search outcomes of UI claimants. Some of them can take the form of rapid-cycle tests of information interventions, whereas some may be more aligned with an earlier generation of UI experiments that tested broader features of the program (Meyer 1995). Chapter V includes some suggestions of changes to broader program features that could be tested in the future.

The projects described in this report represent one set of early, low-cost, rapid applications of experimentation on work search outcomes in UI that address compliance with program requirements. As discussed in each of the two chapters following this one, the interventions we tested in partnership with Washington and North Carolina primarily targeted behavioral dimensions of UI program compliance, rather than dimensions of work search more broadly. These projects contribute to an emerging evidence base on this more general set of research frontiers indicated by the literature—areas of inquiry that have important implications for the effectiveness and efficiency of the overall UI program.

III. Testing a Work Search-Focused Communication Intervention Informed by Behavioral Science in North Carolina

To build evidence on whether applying insights from behavioral science could improve work search among UI claimants, our team partnered with two states to develop and test behavioral interventions tailored to the unique conditions in each state. This chapter summarizes our work in North Carolina. Appendix A provides additional details on several aspects of the study, including technical details of our quantitative methods.

A. State context

We partnered with North Carolina because addressing work search errors is a high priority for state and agency leadership. In the 2019 Improper Payment Information Act (IPIA) reporting year, the overall overpayment rate was 16 percent. Although this was a substantial improvement over historical trends in the state—from 2015 to 2019, the average overpayment rate was just over 20 percent—it was still higher than the national average rate of 10 percent, as measured by the Benefit Accuracy Measurement program (U.S. DOL 2019). Work search issues were the most common cause, accounting for 46 percent of all overpayments in North Carolina. Our review of historical data trends suggests that an intervention to address the state's work search challenges might meaningfully further reduce overpayments attributable to work search errors.

Our work to develop and test a behavioral intervention covered several aspects of the UI system in the state, including North Carolina's work search requirements, the filing methods claimants use, and features of the state's UI and workforce programs (see Exhibit III.1 below).

<div style="border:1px solid #000; padding:10px;">

Exhibit III.1. North Carolina UI system features (as of July 2018)

- **Work search requirements.** Starting in July 2018, North Carolina required claimants to contact three employers during any week in which they claim benefits. The contacts could all be made on the same day and still meet the work search requirement. For each contact, claimants are required to record the date, name of the employer, contact method, employer contact information, position sought, and results of the contact. Claimants are further instructed to maintain their work search records for at least five years. To document their work search, they may use either a form provided by the state (NCUI 506E) or their own personal documents. The amount of information claimants must understand about both what counts as work search and how work search must be recorded gave us scope for designing an intervention aimed at effectively communicating with claimants about these issues.

- **Claim filing methods.** Claimants primarily file their claims online or over the phone; those filing claims online must provide valid email addresses. This requirement made North Carolina a promising state for testing an intervention delivered through email and/or the online claims portal. In 2018, roughly three-quarters of claimants filed their initial and continued claims online (U.S. DOL 2018). Claimants filing online use the state's Southeast Consortium for Unemployment Benefits Integration (SCUBI) information technology system.

- **UI administration.** Since 2012, the North Carolina Division of Employment Security has administered and verified UI benefits in North Carolina and has contracted with the Division of Workforce Solutions (DWS) for delivery of reemployment services. (Prior to 2012, both units were in the same division.) After the first claim is paid, DES sends a file to DWS to schedule a review within four weeks. The review is either a Reemployment Services and Eligibility Assessment (RESEA) appointment (if the claimant qualifies for RESEA) or an Employability Assessment Interview (EAI). All claimants receive a request to attend one of the two meetings. Because the meetings include a review of claimant work search, they offer an opportunity to observe the incidence of improper payments due to work search (as determined through the state adjudications process).

- **UI duration and amount.** On average, claimants in North Carolina receive benefits for a short time (just less than nine weeks in 2019), and the state has a low weekly maximum benefit amount ($350) relative to other states. One implication of the short duration for receiving benefits is that the window during which the state can interact with claimants to encourage active work search is relatively short. In practical terms, by the time work search issues are identified at in-person RESEA or EAI meetings, claimants may have filed for several weeks of benefits; by that time, it may be harder to influence their work-search behavior.

</div>

In addition to key features of the state's UI system, two other factors shaped the study.

Coordination with other efforts. During the time our team was working with North Carolina, the state was concurrently participating in a parallel effort aimed at reducing UI improper payments. Funded by DOL through the UI Integrity Center under the National Association of State Workforce Agencies (NASWA), the State Intensive Services (SIS) initiative sought to help individual states address the root causes of improper payments through a variety of operational, policy, training, and messaging strategies. One component of the SIS initiative was to offer states support for behavioral intervention design and testing, delivered through the SIS team. As our work in North Carolina proceeded, we coordinated closely with NASWA's SIS team to share information and to consider how each of our efforts might complement the other. One critical area of joint planning was to finalize the sequence of intervention development and testing. Early discussions with DOL, North Carolina, and the SIS team led to a decision that our team would go first and focus on areas conducive to rapid testing to inform the SIS team's thinking. The SIS team would build on DOLBI findings and field its intervention only after our testing had finished. One implication of this approach was that our team needed to work quickly to develop and test an intervention, both to help address state needs in a timely manner and to avoid delaying the SIS team's work. Our coordination with the SIS team also helped in that regard; we were able to quickly access state data by executing a data sharing agreement the SIS team already had in place with the state, and our team joined the SIS team's site visit to streamline our information-gathering process, thus reducing the burden on the state.

COVID-19 pandemic. Our work with North Carolina began in summer 2019. The COVID-19 pandemic began impacting the United States in early 2020 and had important effects on the North Carolina study. In response to deteriorating labor market conditions caused by pandemic-related shutdowns, North Carolina suspended its work search requirements in March 2020. This suspension meant that we had to truncate slightly the sample period for one component of our quantitative analysis. Further, as the pandemic unfolded, the state saw an unprecedented increase in the volume of UI claims. Handling the flood of claims quickly became DES's top priority, which precluded additional collaboration around refining the interventions and pursuing further testing to improve results.

B.   Behavioral audit and study development

Our work with North Carolina began with a review of publicly available information (for example, state websites, manuals, Frequently Asked Questions, work search logs, and other documentation available online), focused on both identifying areas in which the state interacted with claimants and identifying possible behavioral barriers. We used a structured template to guide our review and capture our assessments of the likely effectiveness of each message to claimants, focusing on key dimensions informed by behavioral science:

- **Attracting initial engagement.** Where was relevant information on work search requirements located? Was it easy to find and prominently showcased with compelling subject lines?

- **Making required actions easy to understand and memorable.** Was clear, complete, and consistent information provided regarding (1) conducting work search requirements, (2) maintaining work search logs, and (3) maintaining supporting documentation for these logs?

- **Motivating action.** Were the benefits of action clearly laid out? Were deadlines clearly laid out and emphasized? Did the materials appeal to loss aversion—that is, people's tendency to give more weight to an anticipated loss than they would to a similar-sized gain? Did they leverage social norms—for example, encouraging compliance by noting that a large majority of claimants comply with requirements?

- **Removing obstacles to action.** Was the required response easy to make? Were there unexpected hassle factors or unnecessary extra steps? Was assistance offered?

Our review gave us important insights into the state's requirements and how they were communicated to claimants. The work search requirement of making three contacts was relatively straightforward and clearly communicated. On the other hand, information on how to properly document work search contacts was harder to find and understand. Following the review, our team conducted a three-day in-person site visit to gather additional information. The visit focused on helping us better understand the sequence of steps claimants went through; identifying additional interactions among DES, DWS, and claimants; and learning about the state's administrative system capacities. We spent the on-site time meeting with state staff, including both senior leadership and staff members. Throughout the information-gathering process, we closely coordinated our state engagement with the SIS team to minimize the overall burden on the state and help speed the process.

Our synthesis of the materials review and discussions with state staff suggested a series of behavioral bottlenecks (see Chapter II) that could be contributing to claimants' work search errors. These bottlenecks align with the behavioral science findings related to program compliance that were discussed in Chapter II. Specifically, we identified the following bottlenecks related to the salience of program compliance and ways that communications could make compliance easier for claimants to achieve:

- **Inattention and lack of timely communication.** DES staff believed that claimants did not read the information about work search requirements carefully when filing online. Although the details of the number of required work search contacts are explained simply and clearly on the DES website, the information was nested in an FAQ section that claimants may or may not view. Program and frontline staff indicated their impressions that people navigate through the site quickly.

- **Misunderstanding of work search.** DES staff perceive that claimants do not always understand what counts as a valid work search contact—especially for job searches conducted online. They also perceive that claimants fail to understand that work search must occur every week and/or may not understand how many work search contacts are required. Compliance seemed to improve after the in-person RESEA/EAI meeting, during which workforce staff might point out shortfalls in documentation and refer the case for adjudication.

- **Complexity of maintaining adequate supporting documentation.** DES and DWS staff shared that claimants often do not keep a record of their correspondence with employers, and there is limited guidance on what documentation to keep. Moreover, reporting this information correctly to DES involves additional complexity: Claimants often do not use the form provided for logging work search activities or may use it incorrectly.

- **Lack of understanding of work search review process.** We noted that North Carolina's communications did not emphasize that everyone's work search record would be reviewed in person at either an RESEA or EAI meeting. If claimants were unaware of this requirement, they might doubt whether a verification process exists because North Carolina does not require claimants to submit logs or supporting documentation during the weekly claim certification process. Also, given the short average duration of receiving benefits in North Carolina, the verification steps at the RESEA/EAI meetings typically occur for most claimants only late in their claim period, so claimants may consider the repercussions of poor compliance to be low. If claimants do not believe their work search activities will be checked, the emphasis on keeping records for five years might make some of them overlook the verification that will occur much sooner at the RESEA/EAI meetings.

With these bottlenecks in mind, we worked with staff in North Carolina to quickly prototype an intervention aimed at helping claimants improve their work search by reducing the influence of the bottlenecks. We focused on identifying the most feasible levers to target by using a low-cost intervention that afforded opportunities for rigorous testing.

C. Intervention features

During our in-person visit to the state, we identified three promising intervention opportunities:

1. **Delivering a system alert to claimants filing their weekly claims online.** DES had the ability to set a system alert within the Southeast Consortium for Unemployment Benefits Integration (SCUBI) system that would appear when claimants logged in to file a weekly certification. The alert would appear right after the claimant logged into the system and require the claimant to click on an acknowledgment button before proceeding to the main landing page. Once it was acknowledged, the alert would no longer appear until the system made available the opportunity to claim the next week's benefits.

2. **Sending reminder emails to claimants who had an alert set by DES.** DES could supplement the system alert with an email to claimants. The email could provide the same information contained in the alert; it also would give claimants a way to review the information later.

3. **Providing work search tip sheets with guidance about the types of contacts that count and how to document work search appropriately.** DES was willing to make available online a web page or PDF with tips about work search. This approach offered a way to share additional materials targeting the more complex areas of misunderstanding among claimants. DES was also willing and able to embed hyperlinks to the tip sheets within both system alerts and emails.

Working with DES staff, our team prepared mock-ups of the intervention materials and held a series of meetings with them to explain how different features of the materials addressed the behavioral bottlenecks. We also discussed how the materials would be delivered and collaborated on the final design. We then finalized the logistics of fielding the materials and incorporating evaluation procedures.

Key features of the alert and email included the following:

- Starting with an upbeat motivating message about the personal benefits of early, timely, and regular work search
- Using simple language to clarify the need for three employer contacts per week
- Using bulleting, color, icons, and brief bolded text to make the three necessary steps for work search compliance both salient and memorable
- Reminding claimants of the consequences of incomplete records
- Linking to tips that would help them execute the steps well

Exhibits III.2–III.4 show each of the three intervention components (the content of the system alert and email was nearly identical).

Exhibit III.2. Intervention alert

Starts with an upbeat message emphasizing benefits

Clearly and concisely explains key aspects of work search requirements—uses visual cues

Makes number of required job contacts prominent

Claimants are required to acknowledge the alert before continuing

Exhibit III.3. Intervention email

Starts with an upbeat message emphasizing benefits

Makes number of required job contacts prominent

Makes work search requirements salient by communicating about them separately

Reduces barriers to action by providing direct links to tools

Highlights possible consequences of failure to meet requirements

Amin, Samia

From:	DoNotReplyDEV@nccommerce.com
Sent:	Monday, July 15, 2019 10:04 AM
To:	Amin, Samia
Subject:	Stay on top of your job search while claiming UI benefits

Dear job seeker,

Starting early and looking for work every week can really pay off—people who apply for work regularly are more likely to find a job quickly. And you need to make **3 job contacts** each week to keep getting UI.

Your records of work-search efforts will be reviewed, so make sure to follow these 3 steps:

- **Search.** Plan and complete **3 job contacts every week.**
- **Record.** Write down key details about your weekly work search using this form.
- **Save.** Save supporting info (such as emails sent) to provide proof when asked.

If your records are incomplete, you may have to pay back your UI benefits.

Need help figuring out what to do? Use these tips to get it right.

Sincerely,

Division of Employment Security
Customer Call Center
P.O. Box 27967
Raleigh, NC 27611
Tele: 1-888-737-0259
Fax: 919-250-4315
Hours: 8:00 AM to 5:00 PM Monday-Friday

Exhibit III.4. Intervention tip sheet

Callout boxes around the tip sheet:

- Clarifies guidance about what counts as work search and what does not
- Provides a checklist for keeping track of tasks
- Makes key information easy to scan by organizing it by work search contact method
- Gives brief step-by-step instructions in simple language

D. Evaluation design

A key part of our work in North Carolina was to embed a rigorous evaluation to maximize learning. Our evaluation design focused on answering the following key research questions about how the intervention was implemented and its impacts:

Implementation questions

1. What did we learn about our behavioral diagnosis and intervention? What were staff and claimant perspectives on behavioral barriers?
2. Was the intervention applied as intended, and how did staff and UI claimants perceive it?
3. How did the context affect the study and our interpretation?
4. Did claimants see the alert, email notifications, and tip sheets?

Impact questions

5. Did the intervention change claimants' knowledge of work search requirements?
6. Did the intervention reduce claimants' rates of work search errors and overpayments?

When considering design options for the impact study, we sought to balance its rigor with the state's sense of urgency in trying as many potentially useful strategies as possible on a large scale. DES was eager to send the work search system alert and email to all claimants in the state, so one goal of the design stage was to accommodate that preference by keeping the experimental portion of the evaluation as short as possible.

In addition, our design options had two key data-related constraints. First, we relied on a proxy measure of claimants' knowledge of work search requirements. Specifically, we used administrative records indicating, for each weekly UI claim, whether a claimant self-attested that they had completed the required work search activities that week. If claimants indicated on a weekly claim that they had not completed their work search requirements, they were coded as *not* having knowledge of the work search requirement. This metric assumes that if those claimants who attest to noncompliance had fully understood that work search was a requirement for receiving benefits, they would not be filing a claim. We recognize that this assumption is an imperfect one—for example, some proportion of claimants may have known about the work search requirements all along but did not believe they would be enforced, or they may have believed they were supposed to file a claim regardless of noncompliance and let the UI agency confirm their ineligibility. Second, limited Benefits Accuracy Measurement (BAM) sample sizes and the low incidence of work search improper payments detected by North Carolina's adjudication system meant that an analysis based on those outcomes would require a larger sample (that is, a longer duration), which had to be balanced against the state's desire for a shorter testing period.[3]

Our final evaluation design consisted of four study components organized into two phases:

1. **Phase 1:** (1) A randomized controlled trial (RCT) focused on measuring the impacts of the intervention on claimant knowledge of work search requirements and (2) an analysis of data on claimants' engagement with intervention components. For the RCT, we randomized claimants with 50 percent probability to a treatment group that received the intervention or a control group that did not, based on a randomly generated claimant identification number. This phase used data for a sample of 24,416 claimants who filed initial claims in a six-week period during July and August 2019.

2. **Phase 2:** (3) A quasi-experimental design (QED) pre-post analysis estimated the impacts of the intervention on work search behavior, as measured by state adjudication data on work search errors and an analysis of the likelihood that an individual work search contact was deemed acceptable, using BAM data. This phase used data for claimants (183,724 claimants for the adjudications analysis and 1,455 work search contacts from 461 claims for the work search contact acceptability analysis) who filed claims from November 2018 to February 2020. The intervention was rolled out to all claimants in October 2019, marking the beginning of the design's "post" period.

3. **Both phases:** (4) Qualitative research enhanced our understanding of stakeholder perspectives. This phase included data collected from interviews with state staff and focus groups with claimants.

4. **Iterative learning.** One objective of the evaluation was to support an iterative learning process with the state to understand the effectiveness of the initial intervention, learn about ways it might be improved, and then test a more refined intervention.

Our hybrid learning approach, featuring both an RCT and a QED pre-post analysis, was driven by North Carolina's eagerness to field intervention components to all participants, as well as its plan to further partner with NASWA's SIS team on additional behavioral interventions. The state agreed to field the intervention selectively for a limited amount of time (July–August 2019), followed by fielding it for all claimants beginning in October 2019. This approach both gave us a sufficient sample for the RCT to measure proximal outcomes and allowed us to measure outcomes requiring a larger sample size via the

---

[3] The BAM program is a national UI monitoring program that, among other activities, selects a random sample of paid UI claims for states to audit so they can determine whether the payments were proper. For most states, BAM samples roughly 40 paid claims per state per month. The review process involves, in part, state staff attempting to verify work search information submitted by claimants. BAM is the authoritative source of data on state improper payment rates and the causes of improper payments (for example, work search).

QED analysis. However, due to the COVID-19 pandemic, many states, including North Carolina, temporarily waived work search requirements in spring 2020 when the UI system began experiencing record-high numbers of claims. Because of this, the available sample for the QED analysis was slightly truncated, reducing our post-period sample window by one month. Additionally, the ongoing health crisis made it infeasible to design and test further refinements to the intervention. Appendix A provides more details on each of the four study components, the outcomes of interest, data sources, and our analysis methods.

E. Key findings

The study in North Carolina revealed a number of important findings that represent the first step of a longer-term learning agenda to use behavioral science in reducing work search improper payments. We have organized our findings into two groups: those from the implementation analysis and those from the impact analysis. In Chapter V, we discuss the broader implications of our findings—from both North Carolina and Washington—for future work in developing behavioral intervention studies and thinking about how to build on our work to continue testing ways behavioral interventions might improve work search.

1. Implementation analysis

Qualitative data on implementation and study context. To understand the perspectives of state staff on the intervention materials and their experience in collaborating on the study, we conducted brief semi-structured discussions with two senior DES staff in fall 2019 (soon after the RCT intake ended). The discussions focused on four main topics: (1) the level of effort and resources involved in the design and implementation of the intervention, (2) the degree to which the intervention was implemented as intended, (3) key features of the context in which the intervention was implemented, and (4) lessons learned through the design and implementation processes. From our conversations with state staff, we learned the following:

- **State leadership found the intervention materials to be responsive to what they perceived as the causes of work search errors.** In their view, misunderstanding of when to begin work search, how to document it, and how long to retain records were the main drivers of work search errors.

- **The state scaled the use of the intervention materials. The state used the intervention materials in a broader effort to reduce work search errors.** The state incorporated posters displaying the materials into training sessions for DWS staff, programmed the alert and sent the reminder emails to all claimants, and displayed large posters of the tip sheet at American Job Centers.

- **Leadership in DES was interested in exploring the application of behavioral insights to other interactions with claimants, through both the UI office and DWS. DES reached out to the study team to brainstorm additional ways to apply insights from behavioral science to improve operations.** Leadership was especially interested in making BI-informed revisions to its work search log and integrating materials from the intervention into invitations to the RESEA program. (This program offers claimants identified as likely to exhaust UI benefits, as well as other UI claimants the state might consider a priority, more intensive work search supports that are typically delivered in person. North Carolina uses the first RESEA appointment as an opportunity to verify work search logs and supporting documentation and sends out those requests as part of the RESEA invitation). DES leadership was also open to working with DWS to see whether behavioral insights could be used to improve the RESEA program experience itself.

Insights from claimant focus groups: In addition to speaking with state staff, we conducted a series of in-person focus groups with a total of 20 claimants—who would have been exposed to the intervention via the state's broader rollout of the alert and email—to understand their perspectives on work search and the intervention materials. (The focus groups consisted of a convenience sample of RESEA participants attending appointments at the Capital Area NCWorks Career Center in Raleigh, North Carolina, on February 4–5, 2020. Each focus group included 2 to 4 individuals). The study team transcribed and analyzed the focus group discussion data using a coding scheme to identify themes and areas of variation. The coding system included categories related to awareness of and feedback on intervention materials, knowledge of work search requirements, perceptions of ease of adherence to work search processes, determinants of work search, and suggestions for process improvements.

From our discussions with claimants regarding their perspectives on work search, we learned the following:

- **Claimants understood the number of required work search contacts.** Those with whom we spoke indicated they were making many more than three work search contacts per week and did not need a reminder about the required number. This suggests that the claimants already had clarity regarding the number of work search requirements needed that week.

- **Claimants were less clear about documentation and record retention requirements.** Claimants used a range of documentation methods, including the form provided by the state or their own electronic or paper records. Some of them knew they needed to keep records for five years. A few claimants shared a perception that the information presented in the invitation letter from DWS for the RESEA appointment did not align with the requirements posted online regarding work search documentation.

- **Claimants indicated that the log templates provided by the state were cumbersome and did not aid their work search efforts.** Several claimants noted that a work search log was not helpful to planning and executing work search and maintaining it was a distraction. They also noted challenges with the format of the work search log, such as its not being editable online, and the form not providing enough space for entering the level of detail required.

- **Some claimants noted the effects of communication tone on morale and motivation for work search.** They explained that it was difficult to maintain an optimistic outlook and continue work search despite being unsuccessful in finding a job. Some claimants indicated that the occasionally harsh tone of official communications from the state exacerbated this challenge.

- **Some claimants referenced structural elements that affect work search.** Some claimants told us that the perceived low level of benefits (North Carolina's maximum weekly benefit amount is $350) gave them a sense of urgency regarding work search.

- **Claimants also made observations about the broader continuum of interactions around work search.** Several claimants were unclear about the purpose of the RESEA meeting. Most perceived it as a process for verifying work search efforts; several respondents were concerned that it might be used to suspend their access to benefits. Only a few of them displayed eagerness about gaining access to work search supports. Some expressed concerns about the way the meetings were scheduled (required attendance at a pre-specified time) and the format (in person versus virtual), both of which created logistical challenges. Exhibit V.1 in Chapter V provides more detailed recommendations on potential areas of improvement for the RESEA program using insights from behavioral science.

From our discussions with claimants about their perspectives on the intervention, we learned the following:

- **Claimants recalled seeing the system alert, but said they only skimmed it.** Virtually all of the claimants with whom we spoke said they saw the system alert when they logged in, but most of them appeared not to have read the information in the alert very closely. None had noticed or clicked on the embedded link to the tip sheet.

- **Few claimants recalled seeing the email.** Several claimants said they received a lot of email communications from the UI office and DWS—including survey invitations—so they tended not to pay much attention to messages from those senders.

- **Virtually no claimants had accessed the tip sheet, though they thought it looked useful.** At most, only a couple of the claimants may have seen the tip sheet. When given a copy to review, claimants preferred the format and content of the tip sheet over other instructions they had received.

- **Claimants felt that seeing the tip sheet earlier would be useful.** After reviewing the tip sheet, some of the claimants suggested that receiving a copy of it early, such as during the week before benefits started, would have been useful.

Engagement findings. To assess claimants' engagement with the system alert and intervention email, we analyzed data on claimants' interactions with the alert, whether they opened the email, and whether they clicked on an email link to view the online tip sheet. Our analysis of the engagement data showed the following:

- **A majority of UI claimants acknowledged the system alert—the main intervention.** Using data from North Carolina's SCUBI online claims filing system, we observed that, of the treatment group members who had the system alert set for them, 78 percent acknowledged it.[4]

- **There was low engagement with the online tip sheet embedded in a link in the alert and email.** Claimants were exposed to the tip sheet only if they clicked on links in the alert or the email. No more than 22 percent of the treatment group who received the alert and email (and possibly 15 percent or fewer) viewed the tip sheet.[5] These low levels of engagement with the tip sheet are consistent with what we heard from focus group participants—that few recalled seeing the tip sheet.

- **The emails did not result in repeat use of the tip sheet as intended.** Only a small fraction (16 percent) of claimants who clicked the email link to visit the tip sheet did so more than once.

---

[4] According to data from DES's platform, 20 percent of the treatment group did not receive the alert and email because of a technical issue—the software program the state ran to set the alert (which subsequently determined whether the claimant was identified to receive the work search email) inadvertently missed these claimants because of the timing of when the software program was run. Our approach to estimating impacts accounted for this issue. Although this did not result in a decrease in the number of claimants used in the analysis, it still led to a decrease in statistical power. The actual minimum detectable impacts were approximately 25 percent larger than they would have been if the alert had been set for everyone in the treatment group (see Bloom 2006).

[5] The data on visits to the online tip sheet did not allow us to identify the specific claimants who viewed the tip sheet, meaning we could not observe whether some visits originating from the email link and the alert link were in fact visits by the same claimant. The percentages reflect different assumptions about the degree of overlap between the group of claimants that accessed the tip sheet using the email link and the group that accessed it using the alert link. Appendix section A.E provides additional details.

2.   Impact analysis

We carried out two separate quantitative analyses that focused on understanding the impacts of the intervention on outcomes related to work search: the RCT analysis and the QED analysis. Below, we highlight the main results from each analysis.

RCT findings. When claimants file their weekly benefit claim, they are asked to indicate whether they completed the required work search that week. Our key outcome measure for the RCT was an indicator for whether claimants self-attested that they did not complete work search. Our RCT analysis found no impact on this measure of claimant knowledge of work search requirements. Exhibit III.5 shows the rate at which claimants in both the treatment and control groups asserted that they did not conduct work search, based on administrative data from the SCUBI system. We found that the treatment group was less likely to self-attest that they did not conduct work search by two-tenths of a percentage point, but the difference was not statistically significant. The lack of a statistically significant finding may not be surprising given the low engagement with the tip sheet and email discussed above. On the other hand, since claimants in the treatment group did receive and acknowledge the intervention alert, the engagement data does not clearly suggest that this finding was to be expected.

Exhibit III.5. Impacts on claimant knowledge of work search requirements
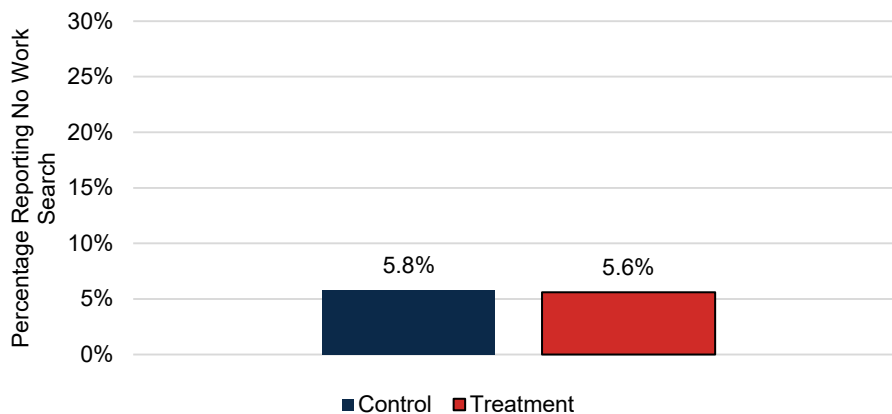
Source: SCUBI data.

Note: The control group mean represents the unadjusted proportion of claimants who ever reported that they did not conduct work search, excluding the first claimed benefit week. The treatment group mean represents the control group mean plus the estimated impact from an instrumental variables regression, a technique we use to estimate impacts among those in the treatment group exposed to the intervention (see Appendix A for technical details). The difference between groups was not statistically significant.

QED findings. Our QED pre-post analysis found no impact of the intervention on either overpayments, as measured using state administrative data on adjudications, or the acceptability of work search contacts, as measured using BAM microdata on individual work search contacts. Exhibits III.6 and III.7 show the estimated impacts on these two outcomes.

We found that the intervention was associated with a less than one-tenth of a percentage point reduction in the likelihood that a claimant had an overpayment within the first five weeks; the difference was not statistically significant (Exhibit III.6).

Exhibit III.6. Impacts on work search overpayments
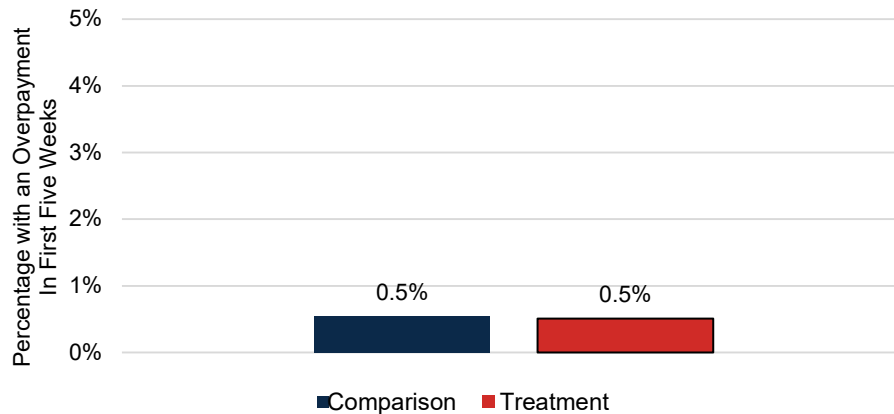
Source: State adjudications data.

Note: Results represent the regression-adjusted impact of making an overpayment within 5 weeks of the claim file date. Results are limited to overpayments determined within 20 weeks of occurrence. We considered all weeks that occurred during the treatment period, or for which a previous week within the same claim had been in the treatment period. The difference between groups was not statistically significant.

We found that the intervention was associated with a 6.4 percentage point increase in the likelihood that an individual work search contact audited by BAM was deemed acceptable (Exhibit III.7). The difference was marginally statistically significant at the 10 percent level, but not at the 5 percent level. This suggests that, while we cannot say with confidence that the intervention increased the likelihood that a work search contact was acceptable, the positive direction of the impact is promising. It may point to the value of additional study to determine more conclusively whether such approaches can improve work search behavior.

Exhibit III.7. Impacts on the acceptability of work search contacts
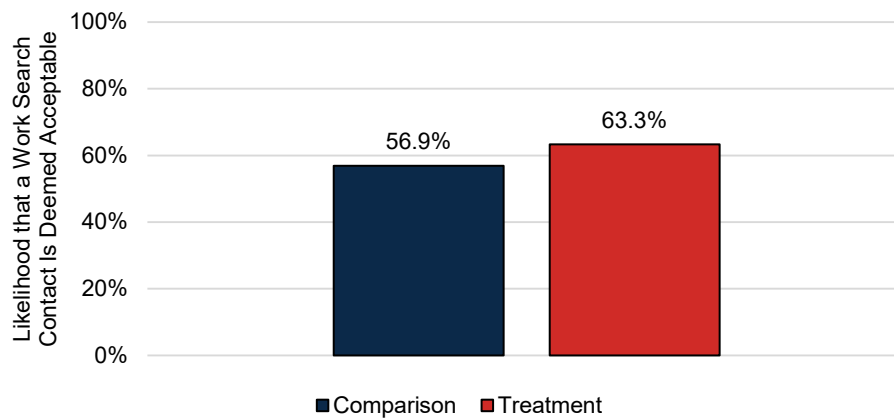
Source: BAM microdata.

Note: * Denotes a program group rate that is significantly different from the estimated control group rate at the 0.10 level. The comparison group value is the weighted proportion of work search contacts in the comparison group counted as acceptable by the BAM auditor. The treatment group proportion represents the control group proportion plus the estimated impact from an inverse probability weighted regression that adjusts for differences in baseline characteristics between the treatment and comparison groups.

In addition to estimating overall impacts for the two outcomes that were the focus of the QED analysis, we also conducted exploratory analyses in which we estimated impacts separately for subgroups defined by (1) having a bachelor's degree or higher, (2) being age 40 or older, and (3) having base period wages higher than the sample median. We hypothesized that the impacts could differ by claimants' level of education, which could influence their understanding of work search requirements in the status quo condition. Similarly, we hypothesized that the intervention might affect claimants younger than age 40 differently than those over age 40, with one group or the other potentially more comfortable interacting with intervention materials delivered through a web-based platform. Finally, base period wages proxies for socioeconomic background, another dimension along which we hypothesized that impacts could vary. We found some evidence of beneficial impacts on both outcomes—the overpayment outcome and the work search contact acceptability outcome—for claimants with at least a bachelor's degree and on the acceptability of work search contacts for those with higher than median base period wages (see Exhibits A.27 and A.28 in Appendix A for details). The subgroup results may suggest promising avenues for further study.

In North Carolina, we were able to quickly develop and test an intervention designed to overcome certain behavioral barriers to successful work search. Our work was carried out within time constraints associated with both the state's interest in rolling out the intervention statewide and DOL's interest in quickly testing an intervention aimed at reducing UI improper payments due to work search errors. Collectively, although our quantitative findings suggest that the intervention had no clear effects on improving outcomes related to work search errors, our qualitative analysis shows that the intervention might have had more positive effects if either a larger share of claimants had read the content of the materials or the materials had been delivered in other ways. Ultimately, our results suggest that, when the conditions in North Carolina (or other states) induced by the COVID-19 health crisis abate, additional refinement and testing of this or similar behavioral interventions may identify promising ways to help claimants reduce work search

errors. Potential directions for future learning, building on the findings from this North Carolina behavioral intervention, are discussed in Chapter V.

IV. Testing a Work Search-Focused Email Informed by Behavioral Science in Washington State

This chapter summarizes our partnership with the state of Washington and the project we conducted there. More background and details of our study design and analysis can be found in Appendix B.

A. State context

Washington's Employment Security Department (ESD), which administers the state's UI system, approached this engagement with a strong interest in testing program innovations designed to reduce improper payments. Like North Carolina, the state has made increasing work search compliance an important priority in its efforts to reduce improper payments, since work search errors have accounted for the majority of overpayments in Washington in recent years (see Appendix Table B.1 for details). These factors made Washington State interested in partnering with the behavioral interventions team to design and test applications of behavioral science to improve work search compliance among UI claimants. Our collaboration with Washington took account of several features of its UI system, including the state's work search requirements and claims filing procedures, in particular (see Exhibit IV.1).

Exhibit IV.1. Washington UI system features as of December, 2019

- **Work search requirements.*** To fulfill their work search requirements, claimants in Washington must contact three employers during each week for which they claim benefits unless they are members of a full-referral union; participate in specially designated training programs; or are exempt for a handful of other, less common reasons. They can substitute approved activities at local workforce development centers for employer contacts. For each contact, claimants must record the date, the name of the employer, the contact method, employer contact information, the position sought, and the results of the contact. Claimants are further instructed to maintain their work search records after they finish receiving benefits. To record the necessary information, they may use either a form provided by the state or their own personal documents.

- **Claims filing procedures.** Because the majority of Washington claimants—67 percent in our analysis sample—file for unemployment benefits through the Internet, email offers a convenient, intuitive way to reach them through the medium they already use to interact with ESD. Washington's system can program emails to be sent to claimants selectively, based on parameters such as the date of filing or the last digit of their Social Security number (our method for selecting claimants to receive the experimental email described below). Also, the online system permits claimant-level linkages of data across systems, which allowed us to study a wide range of claimant outcomes and characteristics.

- **UI duration and amount.** On average, claimants in Washington received benefits for almost 17 weeks in 2019, with a relatively high average weekly benefit amount ($456) relative to other states.

  *These were the work search requirements as of late 2019 and early 2020 when the study was planned and initiated; as described below, requirements changed in March 2020 due to the COVID-19 pandemic.

Impact of COVID-19 on study plans. The study team originally planned to use records from systematic internal review of paid claims as the primary source of outcome data. At the outset of the evaluation, claimants were required to submit job search records if ESD selected them at random for a Job Search Review (JSR).[6] Otherwise, individuals filing their weekly claims generally only had to indicate with a yes or a no whether they had met their work search requirements. For JSRs, ESD randomly selected several hundred claims each week, requested the related work search records, and reviewed them for completeness. For a random 10 percent subset of those reviews, reviewers also attempted to verify the claimant's work search contacts by communicating with the employers listed. Though not identical to the BAM overpayments metric, JSRs provide Washington with a key internal measure of work search-related

---

[6] This requirement changed shortly after the evaluation began, as described below.

overpayments and include a sample of several hundred claims per week, compared to about 500 per year for BAM reviews, making them a valuable outcome measure.[7]

As a result, because of their internal relevance and larger available sample, we initially planned to use JSR-identified work search errors as our primary study outcome. However, the COVID-19 pandemic required a switch to a new source of outcome data because it both shortened the study period by four months and caused ESD to suspend the use of the JSRs as the department responded to the unprecedented volume of UI claims. The new outcome—work search errors identified using ESD's internal adjudication system—used information supplied by claimants when filing their weekly claims to identify errors automatically in addition to reviewer-identified errors.

New work search policy changes. The underlying rate of system-identified errors may have shifted mid-study due to a second change to the evaluation context: Beginning in January 2020, ESD adopted a new policy in which their online filing system began requiring claimants to record the details of their three employer contacts each week or else indicate that they had not completed the requisite work search. Claimants filing by phone were not subject to this new requirement. More details on this policy and COVID-related changes to the evaluation context are described below and in Appendix B.

B. Behavioral audit and study development

In preparation for our coordinated site visit with NASWA's SIS team, who were also engaging with Washington to analyze and develop strategies to address the causes of improper payment issues, we began our engagement with ESD by systematically reviewing publicly available information from Washington's website to identify how the state appeared to interact with UI claimants. We also coordinated with NASWA's SIS team to request and review program information from Washington ESD staff. As in North Carolina, we used a structured template in Washington to guide and capture our assessments along four key dimensions informed by behavioral science:

1. **Attracting initial engagement.** Where was relevant information on work search requirements located? Was it easy to find and prominently showcased with compelling subject lines?

2. **Making required actions easy to understand and memorable.** Was clear, complete, and consistent information provided regarding (1) conducting work search requirements, (2) maintaining work search logs, and (3) maintaining supporting documentation for these logs?

3. **Motivating action.** Were the benefits of action clearly laid out? Were deadlines clearly laid out and emphasized? Did the materials appeal to loss aversion—that is, people's tendency to give more weight to an anticipated loss than they would to a similar-sized gain? Did they leverage social norms—for example, encouraging compliance by noting that a large majority of claimants comply with requirements?

4. **Removing obstacles to action.** Was the required response easy to make? Were there unexpected hassle factors or unnecessary extra steps? Was assistance offered?

We followed our review with a three-day site visit with staff from Washington ESD in August 2019, during which we studied claimant flow through the filing system and identified additional points of interaction and system capabilities. During the site visit, representatives from the study team, in

---

[7] The study team initially considered BAM data as a source of outcome measures. However, given the slow rate of BAM case accumulation—roughly 40 per month—our pre-intervention power analysis determined that even in the most optimistic scenario for using BAM data, using the JSR as our outcome measure would allow for roughly three times greater precision in our impact estimates, while also yielding results several months earlier.

collaboration with colleagues from the NASWA UI Integrity Center, met with 22 state officials (5 senior leaders and 17 staff members) to learn about their perspectives on historical and current drivers of work search requirements and claimant behavior.

Our synthesis of the material review and discussions with state staff suggested the following three behavioral bottlenecks, all of which draw on behavioral science findings related to program compliance:

1. **Lack of timely communication and attention to work search requirements.** Discussions with state staff indicated that many claimants who commit work search errors may do so because the work search requirements are not communicated to them in a clear, salient, and timely manner. One of the primary sources of information on work search, the Unemployed Worker Handbook, presents work search requirements information near the middle of its 74 pages, where it might be difficult to find. Even shorter communications that claimants receive after first applying for unemployment benefits tend to present information about work search requirements near the bottom of a message that contains a lot of other information. As a result, in many cases the BAM investigation or other contacts by ESD staff to verify whether compliance has occurred might be the first time that claimants learn about work search requirements in detail.

2. **Misunderstanding of work search requirements.** According to state staff, when asked to report their work search activities, claimants often describe actions that do not qualify, such as networking, reviewing opportunities on social media, or browsing other job postings online. It is also common for claimants to list activities they completed outside of the reference week that investigators or other staff inquired about. Sometimes claimants also express confusion about whether an activity they completed at a local workforce development office counts toward the required activities. Finally, they report confusion regarding "standby" status—a work search exemption for workers whose employer plans to rehire them. Even though Washington policy requires standby to be verified and reported back to the claimant before the work search waiver goes into effect, some workers said they believed work search requirements were waived simply because they reported being on standby.

3. **Complexity of maintaining adequate supporting documentation.** Claimants often do not have adequate documentation, either because they did not record the correct information about an employer contact or because they did not retain the right information—for example, they may have deleted an application confirmation email when cleaning out their inbox. Claimants also may not know where to find the work search log that ESD created to assist them in reporting work search information.

C. Intervention features

The intervention for this study was an email sent shortly after a claimant's first filed weekly claim, which succinctly explained how to meet work search requirements and linked recipients to additional information. The email served as a high-level reminder of claimants' weekly work search and record-keeping responsibilities, and provided links to Washington's online claim submission website and an example work search log. The email also linked recipients to the section of the claimant handbook on work search requirements, including what is required, who is subject to these requirements, and what claimants need to do to comply with reporting requirements.

An example of the email intervention sent to treatment group members is in Exhibit IV.2. This email incorporated several behavioral elements:

- Making work search requirements salient by communicating about them apart from other broader and more general communications about the UI claims process

- Providing clear information about the work search requirements in a timely and actionable manner at the beginning of the claim

- Reducing information frictions by actively communicating the work search requirements clearly and concisely to new claimants

- Reducing barriers to action by linking directly to online reporting and record-keeping tools, and examples of conforming record-keeping practices

- Highlighting the potential consequences of not meeting the work search requirements

Exhibit IV.2. Example of email sent to treatment group

From: DoNotReply@TBD.com <DoNotReply@TBD.com>
Sent: Tuesday, September 10, 2019 1:41 PM
To: CLAIMANT@YAHOO.COM
Subject: Your job-search requirements while receiving unemployment benefits

[Makes work search requirements salient by communicating about them separately]

Dear job seeker,

Remember: While you claim unemployment benefits, you must look for work every week.

[Provides timely, actionable information]

Here's how to make sure you do what's required:

[Clearly and concisely explains work search requirements]

1. Search: **Complete at least three job search activities every week**. *The only activities that count are:*

- **Contacting an employer about a job or applying for a job.** You can make contacts online, in person, by phone, fax, or mail.
- **In-person activities at a WorkSource office:** These include attending workshops, working with an employment specialist to write a resume, and getting other help. Only approved in-person activities count, so be sure to ask WorkSource staff if the activity qualifies.

[Highlights possible consequences of failure to meet requirements]

2. Record: Save key details about your job search. We may ask to see your job search records. If they aren't complete, we might deny your benefits.

[Reduces barriers to action by providing direct links to tools]

- The easiest way to record the right details is to enter them when you submit your weekly claim online. (You still need to keep a copy of this information for your own records.)
- Or you can use our job search log to record details. If you use this log, be sure to keep your records, even after you receive your last benefit payment.

Need help? Read the section on job search in our Handbook for Unemployed Workers.

Sincerely,

Employment Security Department

D.   Evaluation design

While designing the intervention email in close collaboration with Washington ESD staff, we also designed an evaluation to rigorously test the impact of the email and learn about the process of implementing the email intervention. Our evaluation focused on answering the following key research questions:

Implementation questions

1.   What did we learn about our behavioral diagnosis and intervention?

2.   Did ESD implement the intervention with fidelity to the design? How did staff perceive it?

3.   How did COVID-19 and other external contextual factors affect the study and our interpretation of findings?

4.   Did treatment group claimants open the intervention emails and click through to the informational links?

Impact question

5.   Did the email notification reduce claimants' rate of work search errors and the state's overpayments?

The Washington evaluation used an experimental design to measure the impact of the email intervention on claimant behavior. Using the last digit in their Social Security numbers, we assigned half of the claimants (13,444, or 49.9 percent of the 26,967 eligible study period claimants who provided a valid email address to ESD) to receive the intervention email following their first filed weekly claim; the 13,523 claimants assigned to the control group did not receive the email. The close similarity in characteristics for the treatment and control groups (that is, those who were assigned to receive the email and those who were not), shown in Appendix Exhibit B.3, suggests this assignment mechanism produced similar groups. Treatment group members each received a single email on Tuesday morning in the first week for which they would be eligible to receive benefits. We constructed outcomes using an adjudications database from ESD, which tracked claim-level work search issues and errors, job search review results, and payment activity.

Further details of the evaluation design and construction of outcomes are available in Appendix B, Section C.

E.   Key findings

1. Implementation analysis

Qualitative data on implementation and study context. The study team conducted semi-structured discussions with five key Washington ESD program managers and staff to learn about their experiences in participating in the design and implementation of the email intervention. The discussions gathered information on four primary topics: (1) the level of effort and resources involved in designing and implementing the intervention, (2) the degree to which ESD was able to implement the intervention as intended, (3) key features of the context in which they implemented it, and (4) lessons learned through the design and implementation processes. The study team transcribed and coded discussion data to identify themes and areas of variation. We identified four primary insights from our conversations with ESD staff.

Finalizing the intervention email required commitment, coordination, and effort from Washington ESD teams.

- **Finalizing the intervention required support from agency leadership and identification of key ESD staff members to consult for feedback at important junctures.** Multiple internal teams at ESD play a role in communicating with claimants, and staff we interviewed indicated some of their colleagues had questions about using new, claimant-friendly messaging techniques, instead of the typical approaches to official communications that relied more heavily on direct statutory quotations. Securing buy-in from agency leadership demonstrated to all staff that the intervention was an agency priority, and identifying the appropriate staff to consult on any given question helped shorten the intervention development period. ESD staff recommended carefully considering which staff members and agency leaders should be involved at the start of a new endeavor.

Changes to the online filing system during implementation of the email may have affected the effectiveness of the intervention.

- **In the early weeks of the implementation of the new behavioral intervention, ESD also enacted a new requirement that claimants filing online claims report their job contact information for the week.** Beginning on January 7, 2020, the online system prevented people from completing their weekly claim until they either submitted contact information for at least three employers whom they had contacted for work search or attested they had not completed at least three work search activities.[8] People who submitted their claim by phone were not subject to the same requirement, although this difference likely was not apparent to claimants unless they tried using both systems. After launching this system modification, ESD added functionality that allowed claimants to download their work search entries as an aid to complying with record-keeping requirements.

- **The changes to the online system may have reduced the potential for the intervention email to affect claimant behavior by making it easier (and required) for the roughly two-thirds of claimants who filed online to document work search activities.** The intervention email was designed under conditions in which claimants had to provide evidence of job contacts only if requested—a relatively rare occurrence. However, most of the implementation took place during a period when all online claim filers were required to provide job contacts or else certify they had not completed the required activities. The changes to the system made it more difficult for claimants to file a weekly claim without completing their work search activities and may have lowered the prevalence of work search errors, even in the absence of the email tested in this project. The reporting requirement also made the work search requirement more salient early in the claim process than it had been previously, thereby addressing one of the key behavioral bottlenecks targeted by the email intervention. By increasing the salience of the work search requirement, the online filing requirement may have limited the potential for the informational email to reduce further the prevalence of work search errors among email recipients.

COVID-19 dramatically affected the state's economic context and limited the study's conclusions.

- **COVID-19 led to severe demands on ESD staff and caused the governor to waive work search requirements on March 24, 2020, resulting in a reduced period of intervention implementation and the follow-up period ending four months early.** Washington State was one of the earliest and most serious centers of the COVID-19 pandemic in the United States. On February 29, 2020, the

---

[8] The BI team was aware that this change in requirements was in development, but at the outset of evaluation planning in fall 2019, it appeared most likely that the change would be implemented late in the evaluation period or after the evaluation had ended.

governor's office declared a statewide emergency, followed by 35 additional proclamations related to the virus response in March alone. Many of these proclamations encouraged or mandated measures, such as business closures and stay-at-home orders. Nearly all staff reported that COVID-19 led to rapid, significant increases in the number of UI claims ESD received. The number of UI initial claims filed in the state in the four weeks ending April 4 increased by more than 20 times from the previous four weeks (ETA Form 539); on March 24, 2020, the governor formally waived the work search requirement. By the time the governor announced this waiver, ESD had already responded to the sharp increase in unemployment claims by shifting staff priorities, including suspending JSR activities the week of March 13 to process claims as rapidly as possible. ESD staff reported that the work search waiver led to the early end of the study.

- **These changes limited the availability of outcome data for the study's primary planned outcome measure.** Waiving the work search requirement reduced the sample of UI claimants with data available on the primary outcome of interest—work search errors measured in the JSR data—by almost two-thirds.

ESD's role in reviewing and approving the behaviorally informed email helped prepare staff for rapid adoption of communication changes during the COVID-19 pandemic and focused them on participants' needs in future communications.

- **The development process introduced the mindset of designing communications and support to put claimants' needs first.** The intervention addressed important challenges claimants faced, including receiving clear and timely communications about work search requirements. The experience of reviewing and approving the newly developed intervention helped staff across the agency realize that the agency's messaging was not always as effective as hoped and that it was necessary to change its communication approach. Informed by this experience and by increased feedback from customers during the pandemic, the agency now has a "customer experience team," which speaks directly to claimants about their challenges. Some staff recommended that other agencies adopt a similar approach.

- **The development process unexpectedly helped prepare ESD for the types of processes needed to rapidly develop and approve new language during the COVID-19 pandemic.** Several staff members reported that this experience made it easier to adapt communications in order to formulate a timely response to the developing pandemic.

Engagement with the intervention email. To assess claimants' engagement with the intervention email, we measured whether recipients opened the email and whether they clicked on links within the email. Our analysis of the engagement data showed the following:

- **Most treatment group members viewed the intervention email.** By seven days after the intervention email was sent, aggregate engagement data showed that 69 percent of treatment group members had opened the email.

- **About 5 percent of treatment group members clicked through to the links.** More than 5 percent of claimants had clicked through to the most commonly used link—the weekly claims log-in page. The job search log template and unemployed workers' handbook links received fewer clicks.

Although encompassing a small minority of claimants, this click-through rate is comparable to the rates recently observed in a similar behaviorally informed email intervention.[9]

2. Impact analysis

To assess the impacts of the intervention email, we measured the rates of work search errors and other outcomes among the claimant groups that did and did not receive the email. Specifically, this study measured impacts on three additional outcomes: the share of payments that were later disqualified; work search errors among claims selected for JSR (the originally planned primary outcome before COVID-19); and work search issues that were later cleared, potentially an indicator of claimant confusion. (More information is available in Appendix B.)

Impact on work search errors and other outcomes. To measure the intervention's impacts on work search error rates and other outcomes, we compared the prevalence of each outcome in the treatment and control groups. Because of our experimental study design, the difference between the rate of each outcome in the control and treatment groups—shown below in Exhibits IV.3 and IV.4—is an estimate of the effect of the treatment email on the likelihood of each outcome.

When considering these results, it is worth noting the context described throughout this report. For instance, starting in January 2020, the online claims filing system added a requirement that claimants record their employer contacts before filing their claim or acknowledge that they had not searched for work. This change could have prevented many of the work search errors among both control and treatment group members that otherwise might have been addressed by the treatment. It also reduced the impact of the intervention relative to a hypothetical scenario where it was fielded without this policy change. Similarly, the COVID-19 pandemic reduced the size of the study sample and shortened the follow-up period; this prevented us from using our preferred outcome (that is, work search errors detected by the JSR). Although the number of claimants was still robust, the follow-up period for most of them was substantially shorter than it would have been in the absence of the pandemic, and some claimants had as little as one benefit week in which to respond to the treatment. This again reduced the opportunity for us to observe any difference that might have emerged between the treatment and control groups as a result of the intervention, given more time and a larger sample. More detail on these factors is included in Appendix B, Section A.

---

[9] Executive Office of the President of the United States. National Science and Technology Council. "Social and Behavioral Sciences Team Annual Report, September 2015." Behaviorally informed emails from the Department of Veterans Affairs encouraging veterans to apply for benefits garnered a 43 percent open rate (versus 40 percent for control emails) and a click-through rate on application links of 4.1 percent (versus 3.6 percent for control emails). Available at https://web.archive.org/web/20170217164040/https:/sbst.gov/download/2015%20SBST%20Annual%20Report.pdf. Accessed March 2, 2021.

We found no evidence that the intervention affected the likelihood of any of our outcomes of interest. As shown in Exhibit IV.3, the rate of work search errors was virtually identical in the treatment and control groups, with 6.2 percent of claimants in each group committing an error after entering the study sample.

Exhibit IV.3. Impact on work search errors
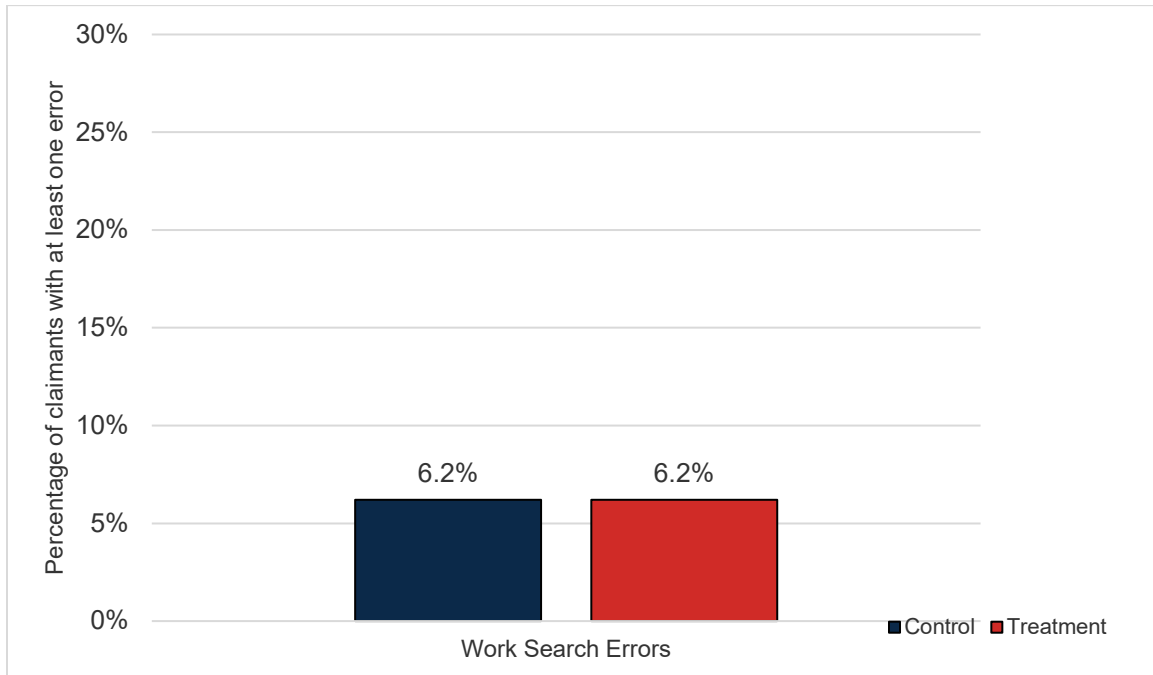
Source: ESD claims adjudication data.

Note: Group means are adjusted for county, pre-separation job, waiting week, and Spanish language preference. The difference between groups was not statistically significant.

In addition to work search errors, we examined three other related outcomes: (1) payments made to claimants that were subsequently disqualified (similar to an overpayment), (2) work search issues "cleared" after a determination that no error had been made, and (3) errors committed by claimants selected for a JSR. Although these outcomes showed some difference in likelihood across the treatment and control groups, none of the differences measured was statistically significant. Moreover, as shown in Exhibit IV.4, the differences were all quite small: The largest difference—on JSR work search errors— was only 1.5 percentage points and 12 percent of the control group rate, and the others were substantially smaller. Additional detail on our analytical methods and findings, and supplementary robustness checks, is available in Appendix B.

Exhibit IV.4. Impacts on error-related outcomes
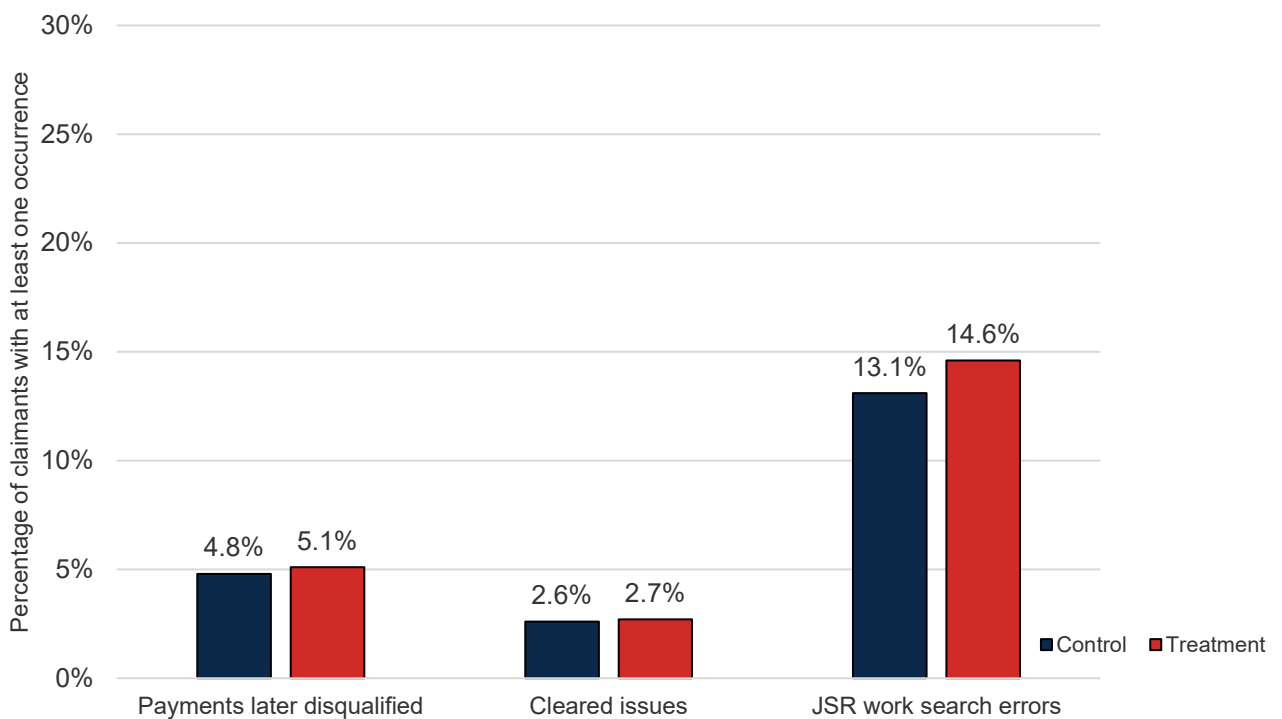
Source: ESD claims adjudication data.

Note: Group means are adjusted for county, pre-separation job, waiting week, and Spanish language preference. The sample for work search errors in the JSR sample comprises individuals selected for a JSR. The differences between groups were not statistically significant.

V. Takeaways and Directions for Future Learning

The projects the DOLBI team undertook with North Carolina and Washington State to address these states' priorities in improving work search behaviors among UI claimants add new insights to the evidence base discussed in Chapter II about the possibilities of, and contextual determinants of, applications of BI in UI programs. They also provide useful lessons on promising practices when developing future behavioral interventions to test to improve UI outcomes, and further directions to explore once states have moved beyond dealing with the COVID-19 pandemic and its aftermath. We briefly discuss key findings and lessons for future work below.

A. Lessons learned about impacts and implementation across both states

- **Quantitative findings suggest limited impacts.** In both states, state administrative data show no impacts on knowledge of work search requirements, as measured through self-attestation of work search or work search behavior. A pre-post analysis in North Carolina suggests some improvement in work search behavior (but statistically significant only at the 10 percent level). In both states, a majority of participants were exposed to the main intervention, but only a small share clicked on the embedded links. This finding had particularly important implications for North Carolina, where an important component of the intervention (a tip sheet) was accessible only via hyperlinks. As noted previously, context (specifically, the effect of the COVID-19 pandemic in both states and concurrent policy changes in Washington State) limited what we were able to learn in the studies of these interventions. These developments prevented the iterative design and further testing that is typical in these types of trials and that we had anticipated would allow us to refine the interventions over time.

- **Qualitative findings on implementation suggest work search may nonetheless be a promising area for continued testing of behavioral insights.** User focus groups in North Carolina (1) supported some initial hypotheses about relevant behavioral barriers (low understanding of how to log work search efforts and requirements for maintaining supporting documentation), (2) negated other hypotheses (there was no confusion about the number of work search contacts required), and (3) suggested additional, more complex hypotheses that had not previously been the focus of our interventions (the difficulty of maintaining morale, the effect of concurrent communications, the timing and format of intervention, and the degree to which poor design of work search documentation requirements actually detracted from work search). It is worth noting that focus group participants were those who showed up at their RESEA appointments, and their understanding and perceptions may differ from those who missed their appointments. Further user interviews, experimentation, and learning could produce valuable evidence on how to address both the initially hypothesized barriers our interventions targeted in North Carolina or Washington and those ideas that were not originally a focus of the behavioral science applications we tested.

- **State partners appear to have an appetite for broader applications of behavioral insights.** Staff in both states found the collaborative design process useful. They mentioned instances of applying behavioral principles and the focus on claimant experience used in the design process to other aspects of their work. In North Carolina, state partners quickly scaled the use of intervention materials, sharing them with all new claimants in the state and exhibiting them prominently in one of their larger American Job Centers. They were receptive to discussions on further iterating materials to try to improve results, and they requested further assistance in applying behavioral insights to improving other program communications and operations (including work search logs and RESEA program outreach). In Washington, partners' staff credited their work on the project with increasing their ability to quickly revise communications going forward.

B. Lessons learned on developing behavioral interventions

These early explorations of applying behavioral insights to improve work search have generated some valuable lessons:

It is feasible to quickly deploy interventions embedded in work search processes. Increased reliance on technology for the UI claims process and claimant communications (websites, emails, robocalls, and text messaging) over the last 15 years provides opportunities for integrating behavioral interventions at a low cost. Many of these systems allow for customization and targeted deployment of messaging and make data accessible, which makes rigorous evaluation feasible.

Data systems are not enough. Having project champions and access to skilled staff—particularly those with expertise in data—matters a lot. Although the underlying data infrastructure matters, the speed with which system features can be leveraged to incorporate behavioral insights ultimately depends on state capacity; the availability of skilled staff; and agency leaders who operate as project champions and ensure a commitment of time and resources, irrespective of other priorities. In North Carolina, where both leadership commitment to reducing work search errors and staff capacity were high, and the required partners (program manager, data specialist, and software vendor) were already collaborating closely on other efforts, the intervention was launched within a month of initiating discussions with the state. Senior leadership and managerial commitment proved valuable for securing the sustained engagement and back and forth needed for the research team to fully understand the strengths and limitations of existing data provided by the state and to monitor implementation.

In Washington, coordination among a larger group of partners and adapting existing administrative systems required more time, but the intervention was still launched rapidly—within four months of initial discussions with the state. The shift to a virtual environment due to the COVID-19 pandemic is likely to provide even more opportunities for integrating behavioral insights. At the same time, the pandemic has led to unprecedented demands on UI program staff and leaders, as record numbers of workers apply for benefits and program staff rapidly implement program changes required by federal pandemic legislation. In the near term, these demands on UI program resources may constrain agencies' capacity to add new priorities, such as exploring applications of behavioral insights.

Site visits help establish understanding and trust, and can accelerate collaboration and understanding of context. In both North Carolina and Washington, a two-day site visit very quickly allowed us to understand the context, the spectrum of opportunities for intervention design and data collection, and the key players and their roles. The visit also allowed us to quickly forge the relationships and support we needed to implement the intervention (launching it within a month of the visit in North Carolina and within four months in Washington) and gain access to the data we needed to test it rigorously. In North Carolina, we created and iterated on an initial prototype version of the intervention alert with staff and leadership input while on site, thus generating enthusiasm and momentum among state partners.

State administrative data can be useful but require some investment to assess their potential. In both studies, we were able to access state administrative data to understand work search knowledge and behavior. To determine data quality and the benefits and limits of these data, however, we did need to spend time with state staff to understand the underlying program processes for data collection and validation.

Early engagement with users at the diagnosis and design stages may improve intervention design. In North Carolina, we had the opportunity to conduct user interviews as part of the implementation study. (In Washington, the COVID-19 pandemic prevented data collection from claimants.) The North Carolina discussions produced insights into behavioral barriers and user contexts that had not surfaced previously in discussions with frontline staff. Although program staff (particularly frontline staff) provide important perspectives, they are not an adequate proxy for understanding user perspectives and experiences. Investing in gathering user perspectives early and at other periodic, strategic junctures may be important. It is likely to provide the opportunity to assess whether assumptions and hypotheses are, in fact, correct. It also allows for early input on behavioral intervention materials and rapid iteration, which are likely to yield a better design.

Context is important, and unanticipated developments can impact testing plans. Despite up-front discussion of risks and close contact with partner sites, our studies in both states were affected by unanticipated events (changes to work search requirements due to the COVID-19 pandemic in both states, and changes in reporting requirements in Washington). These experiences underscore the importance of developing close partnerships and monitoring sites closely.

Aligning different technical assistance and research efforts yields efficiencies. The study team coordinated closely with the NASWA SIS team, which was supporting efforts to lower improper payments in states. Sharing information across the two teams and coordinating site visits and requests for information and data helped reduce state partners' burdens substantially.

C. Implications for the next stage in building evidence

Continuing to test applications of behavioral insights for work search processes, as well as other programs and services offered to UI claimants, seems worthy of further exploration to improve program efficiency and outcomes.

Altering the scale, intensity, mode, timing, and alignment of communications. Our studies suggest that the effect of a single behavioral intervention for UI claimants can be heavily mediated by the frequency and content of other communications from different state agencies, changes in policy, and historical context. For example, in North Carolina, the motivating language and clear messaging included in our intervention materials may have been counteracted by other messaging—such as the tone and content of information on the UI claimant portal and/or website or in emails and letters that the Departments of Employment Security and Workforce Solutions sent. Similarly, because the intervention in Washington only affected email communication, there may be additional opportunities to align other official communication—such as messages that claimants receive when completing weekly claims online—with components of the email message. Stronger applications of information-focused behavioral interventions may involve testing and aligning improvements with a fuller suite of communications that UI claimants and job seekers would receive. These types of improvements may be more potent if coordinated and delivered on a larger scale, such as by multiple agencies that may be communicating concurrently with claimants.

Applying behavioral insights to the design and delivery of reemployment support services may provide insights into how to improve work search more effectively. To date, much of our behavioral testing on work search improvements in the DOLBI project has been through virtual communications from a single agency (emails, alerts). Insights from our site visit in North Carolina suggest that there may

be rich opportunities for applying behavioral insights to multiple touch points along the continuum of UI programs and services provided to UI claimants.

These opportunities include areas that could improve compliance with UI program requirements, as well as aspects that could promote more rapid reemployment through work search. For example, a large number of UI claimants are required to participate in the RESEA program, making them eligible for more intensive engagement with both the UI office and workforce services. Our North Carolina site visit surfaced potential opportunities for applying behavioral insights to improve RESEA services; DES's sustained willingness to brainstorm about how RESEA program processes might be improved using behavioral interventions suggests some appetite at the state level. Exhibit V.1 describes in detail the areas of RESEA where applying behavioral insights could lead to improvements in compliance and employment outcomes. We recognize that insights shared in Exhibit V.1 are based on perspectives from a small sample of participants from a single state and a single site and pertain to topics that go beyond the scope of this specific study; we are nonetheless summarizing these in case they are of interest to stakeholders who wish to pursue broader applications of behavioral insights to UI-related activities.

Focusing on levers other than knowledge and simplification of next steps. Our work search behavioral interventions focused more on improving knowledge and making next steps salient and easy to understand. However, claimant interviews in North Carolina suggest that addressing other mediators of behavior also could be helpful. Interventions targeting these mediators could include addressing dips in motivation and morale in the face of unsuccessful job search and financial stress, or providing planning prompts to facilitate both better work search and work search documentation. Expanding the list of activities that qualify as work search—as was discussed during the Washington site visit and later enacted during the pandemic—is another potentially effective lever (Employment Security Department 2021). This is especially relevant to applications of behavioral insights, if the newly qualifying activities align with widely held social norms pertaining to effective job search in the 21st century, including posting a resume online or registering with a recruiter. Behaviorally informed messaging around the expanded list could boost its effectiveness in promoting compliance and supporting reemployment.

Setting the stage for bolder initiatives. Finally, our interactions suggest that collaborating with states on modest interventions may establish the trust, capacity, and appetite for applying behavioral insights needed for future—and possibly bolder—applications in partnership with these or other sites. Due to practical constraints on both intervention and study planning, applications most likely would not include changes in designs that are dictated by underlying legislation, but that still leaves many possible applications to consider. These applications could include the following:

- Service design improvements mentioned above

- Prototyping different options in advance of a planned policy or program change

- Changes to program features informed directly by the behavioral science evidence on reference dependence mentioned in Chapter II; for example, the evidence on reference dependence suggests possible variations in the optimal path of benefits over a period of UI receipt (DellaVigna et al. 2017; Shavell and Weiss 1979)

- Broader requirements that official program communications use plain language and avoid jargon, as Washington State did (ESSB 5193)[10]

---

[10] Although these applications generally do not include options that require new legislation, the Washington State legislature passed new legislation in 2021 requiring that claimant communications use plain language.

- Experiments to test alternative forms of work search, such as variations in the number of required weekly contacts (Klepinger et al. 1997)

- Experiments to test the effects of varying the intensity of search requirements over a period of UI receipt (DellaVigna et al. 2020) or substituting other activities for work search for some portion of a spell (SBST 2016)

The range of potential design changes that could be tested underscores the value of investing in ongoing collaborative relationships that persist over the medium or longer term. This would allow for multi-phased iterations and testing beyond the initial intervention and trial to increase the odds of success and effectiveness of investments in programs and services. These types of partnerships may have the additional benefit of strengthening the practice of evidence-based decision making among state and local agencies and their capacity for future research-practice partnerships and continuous improvement.

References

Akerlof, George A., and Rachel E. Kranton. "Economics and Identity." *Quarterly Journal of Economics*, vol. 115, no. 3, 2000, pp. 715–753.

Amin, Samia, Greg Chojnacki, Aravind Moorthy, Irma Perez-Johnson, Matt Darling, and Jaclyn Lefkowitz. "Using Behavioral Insights to Increase Retirement Savings: Trial Design and Findings." Washington, DC: Mathematica Policy Research, 2017.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American statistical Association,* vol. 91, no. 434, 1996, pp. 444–455.

Babcock, Linda, William J. Congdon, Lawrence F. Katz, and Sendhil Mullainathan. "Notes on Behavioral Economics and Labor Market Policy." *IZA Journal of Labor Policy*, vol. 1, no. 1, 2012, pp. 1–14.

Baily, Martin Neil. "Some Aspects of Optimal Unemployment Insurance." *Journal of Public Economics*, vol. 10, no. 3, 1978, pp. 379–402.

Becker, Gary. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, vol. 76, no. 2, 1968, pp. 169–217.

Belot, Michèle, Philipp Kircher, and Paul Muller. "Providing Advice to Jobseekers at Low Cost: An Experimental Study on Online Advice." *The Review of Economic Studies*, vol. 86, no. 4, July 2019, pp. 1411–1447.

Black, Dan A., Jeffrey A. Smith, Mark C. Berger, and Brett J. Noel. "Is the Threat of Reemployment Services More Effective Than the Services Themselves? Evidence from Random Assignment in the UI System." *American Economic Review*, vol. 93, no. 4, 2003, pp. 1313–27.

Bloom, Howard S. "The Core Analytics of Randomized Experiments for Social Research." New York, NY: MDRC, 2006.

Card, David, Raj Chetty, and Andrea Weber. "The Spike at Benefit Exhaustion: Leaving the Unemployment System or Starting a New Job?" *American Economic Review*, vol. 97, no. 2, 2007, pp. 113–118.

Chetty, Raj. "Moral Hazard Versus Liquidity and Optimal Unemployment Insurance." *Journal of Political Economy*, vol. 116, no. 2, 2008, pp. 173–234.

Chetty, Raj. "Behavioral Economics and Public Policy: A Pragmatic Perspective." *American Economic Review*, vol. 105, no. 5, 2015, pp. 1–33.

Chodorow-Reich, Gabriel, John Coglianese, and Loukas Karabarbounis. "The Macro Effects of Unemployment Benefit Extensions: A Measurement Error Approach." *Quarterly Journal of Economics*, vol. 134, no. 1, 2019, pp. 227–279.

Chojnacki, Greg, Jonah Deutsch, Samia Amin, Irma Perez-Johnson, Matthew Darling, and Jaclyn Lefkowitz. "Using Behavioral Insights to Help Employers Resolve OSHA Citations: Trial Design and Findings." Washington, DC: Mathematica Policy Research, 2017.

Darling, Matthew, Christopher O'Leary, Irma Perez-Johnson, Jaclyn Lefkowitz, Ken Kline, Randall Eberts, Samia Amin, and Greg Chojnacki. "Using Behavioral Insights to Improve Take-Up of a Reemployment Program." Washington, DC: Mathematica Policy Research, 2017.

Darling, Matthew, Anne Stotler, Amy Johnson, Gregory Chojnacki, Mikia Manley, Samia Amin, and William J. Congdon. "A Literature Scan and Synthesis of Research on Labor-Related Behavioral Science." Washington, DC: Mathematica, 2019.

Darling, Matthew, Jaclyn Lefkowitz, Samia Amin, Irma Perez-Johnson, Gregory Chojnacki, and Mikia Manley. "Practitioner's Playbook for Applying Behavioral Insights to Labor Programs." Washington, DC: Mathematica Policy Research, April 2017.

DellaVigna, Stefano, and M. Daniele Paserman. "Job Search and Impatience." *Journal of Labor Economics*, vol. 23, no. 3, 2005, pp. 527–588.

DellaVigna, Stefano, Attila Lindner, Balázs Reizer, and Johannes F. Schmieder. "Reference-Dependent Job Search: Evidence from Hungary." *The Quarterly Journal of Economics,* vol. 132, no. 4, 2017, pp. 1969–2018.

DellaVigna, Stefano, Jörg Heining, Johannes F. Schmieder, and Simon Trenkle. *Evidence on Job Search Models from a Survey of Unemployed Workers in Germany*, no. w27037. Cambridge, MA: National Bureau of Economic Research, 2020.

Diamond, Peter A. "Wage Determination and Efficiency in Search Equilibrium." *The Review of Economic Studies*, vol. 49, no. 2, 1982, pp. 217–227.

Employment Security Department, Washington State. "Job search requirements." Internet page. Available at https://esd.wa.gov/unemployment/job-search-requirements#list. Accessed 25 July 2021.

Employment Security Department, Washington State. *Unemployed Worker Handbook*. January 2019. Available at https://esdorchardstorage.blob.core.windows.net/esdwa/Default/ESDWAGOV/Unemployment/ESD-Handbook-for-Unemployed-Workers.pdf. Accessed February 12, 2021.

ESSB 5193. Engrossed Substitute Senate Bill 5193. "Concerning unemployment insurance claim adjudicators." Section 3, Chapter 271, Laws of 2021. 67th Legislature, 2021 Regular Session. Washington State Legislature.

Executive Office of the President of the United States. National Science and Technology Council. "Social and Behavioral Sciences Team Annual Report, September 2015." Available at https://web.archive.org/web/20170217164040/https:/sbst.gov/download/2015%20SBST%20Annual%20Report.pdf. Accessed March 2, 2021.

Faberman, R. Jason, and Marianna Kudlyak. "The Intensity of Job Search and Search Duration." *American Economic Journal: Macroeconomics*, vol. 11, no. 3, 2019, pp. 327–357.

Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev. "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." *Journal of Public Economics*, vol. 148, 2017, pp. 14–31.

Karlsson, Niklas, George Loewenstein, and Jane McCafferty. "The Economics of Meaning." *Nordic Journal of Political Economy*, vol. 30, no. 1, 2004, pp. 61–75.

Katz, Lawrence. "Discussion of Alan Krueger's 'Where Have All the Workers Gone.'" Presentation, Brookings Papers on Economic Activity, September 8, 2017.

Kay, Richard. "Proportional Hazard Regression Models and the Analysis of Censored Survival Data." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 26, no. 3, 1977, pp. 227–237.

Klepinger, Daniel H., Terry R. Johnson, Jutta M. Joesch, and Jacob M. Benus. "Evaluation of the Maryland Unemployment Insurance Work Search Distribution." Prepared for the Maryland Department of Labor, 1997.

Knabe, Andrea, and Steffen Ratzel. "Scarring or Scaring? The Psychological Impact of Past Unemployment and Future Unemployment Risk." *Economica*, vol. 78, 2011, pp. 283–293.

Krueger, Alan B., and Andreas I. Mueller. "A Contribution to the Empirics of Reservation Wages." *American Economic Journal: Economic Policy,* vol. 8, no. 1, 2016, pp. 142–179.

Loewenstein, George. "Because It Is There: The Challenge of Mountaineering… for Utility Theory." *Kyklos*, vol. 52, no. 3, 1999, pp. 315–343.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao. "Poverty Impedes Cognitive Function." S*cience,* vol. 341, no. 6149, 2013, pp. 976–980.

Mansournia, Mohammad, and Douglas Altman. "Inverse Probability Weighting." *BMJ,* no. 352, 2016. Available at: https://www.bmj.com/content/352/bmj.i189.full. Accessed July 12, 2021.

Marinescu, Ioana. "The General Equilibrium Impacts of Unemployment Insurance: Evidence from a Large Online Job Board." *Journal of Public Economics*, vol. 150, 2017, pp. 14–29.

Marinescu, Ioana, and Roland Rathelot. "Mismatch Unemployment and the Geography of Job Search." *American Economic Journal: Macroeconomics*, vol. 10, no. 3, 2018, pp. 42–70.

Mazar, Nina, and Dan Ariely. "Dishonesty in Everyday Life and Its Policy Implications." *Journal of Public Policy & Marketing*, vol. 25, 2006, pp. 117–126.

Meyer, Bruce D. "Lessons from the U.S. Unemployment Insurance Experiments." *Journal of Economic Literature*, 1995, vol. 33, no. 1, 1995, pp. 91–131.

Michaelides, Marios, Peter R. Mueser, and Jeffrey A. Smith. "Do Reemployment Programs for the Unemployed Work for Youth? Evidence from the Great Recession in the United States." *Economic Inquiry*, vol. 59, no. 1, 2021, pp. 162–185.

Mortensen, Dale T., and Christopher A. Pissarides. "Job Creation and Job Destruction in the Theory of Unemployment." *The Review of Economic Studies*, vol. 61, no. 3, 1994, pp. 397–415.

Mueller, Andreas I., Johannes Spinnewijn, and Giorgio Topa. "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias." *American Economic Review,* vol. 111, no. 1, 2021, pp. 324–363.

Mullainathan, Sendhil, Joshua Schwartzstein, and William J. Congdon. "A Reduced-Form Approach to Behavioral Public Finance." *Annual Review of Economics*, vol. 4, no. 1, 2012, pp. 511–540.

Mullainathan, Sendhil, and Eldar Shafir. *Scarcity: Why Having Too Little Means So Much*. New York: Macmillan, 2013.

Rothstein, Jesse. "Unemployment Insurance and Job Search in the Great Recession." Brookings Papers on Economic Activity. Washington, DC: Brookings Institution, 2011.

Shu, Lisa L., Nina Mazar, Francesca Gino, Dan Ariely, and Max H. Bazerman. "Signing at the Beginning Makes Ethics Salient and Decreases Dishonest Self-Reports in Comparison to Signing at the End." *Proceedings of the National Academy of Sciences,* vol. 109, no. 38, 2012, pp. 15197–15200.

Social and Behavioral Sciences Team. "Annual Report." Washington, DC: National Science and Technology Council, Executive Office of the President, 2015. Available at: https://web.archive.org/web/20170217164040/https:/sbst.gov/download/2015%20SBST%20Annual%20Report.pdf. Accessed March 2, 2021.

Social and Behavioral Sciences Team. "2016 Annual Report." Washington, DC: National Science and Technology Council, Executive Office of the President, 2016. Available at: https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/images/2016%20Social%20and%20Behavioral%20Sciences%20Team%20Annual%20Report.pdf. Accessed March 16, 2021.

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

Shavell, S., and L. Weiss. "The Optimal Payment of Unemployment Benefits Over Time." *Journal of Political Economy*, vol. 87, no. 6, 1979, pp. 1347–1362.

Spinnewijn, Johannes. "Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs." *Journal of the European Economic Association*, vol. 13, no. 1, 2015, pp. 130–167.

U.S. DOL. "2019 IPIA 1-Year Data." Unemployment Insurance Payment Accuracy Datasets. Available at https://www.dol.gov/general/maps/data. Accessed 9 March 2021.

U.S. DOL "2018 IPIA 1-Year Data." Unemployment Insurance Payment Accuracy Datasets. Available at https://www.dol.gov/general/maps/data. Accessed 9 March 2021.

U.S. DOL "Data Downloads." ETA-9128. (Internet page). Available at https://oui.doleta.gov/unemploy/DataDownloads.asp. Accessed 25 July 2021.

Winkelmann, Liliana, and Rainer Winkelmann. "Happiness and Unemployment: A Panel Data Analysis for Germany." *Applied Economics Quarterly*, vol. 41, no. 4, 1998, pp. 293–307.

APPENDIX A
North Carolina Technical Appendix

This appendix provides additional details on the North Carolina study to supplement the information provided in Chapter III.

A. State context

Improper payments and work search issues. As noted in Section III.A, North Carolina has been motivated to address work search error rates. Exhibit A.1 presents key improper payment data for North Carolina from 2015 to 2019. Work search errors were the leading root cause of overpayments in each of these five years.

Exhibit A.1. Annual selected improper payment outcomes, North Carolina, 2015–2019

Measure	2015	2016	2017	2018	2019
Overpayment rate	16.9%	22.1%	23.1%	23.5%	16.2%
Overpayment responsibility: UI claimant only (percentage of dollars overpaid)	80.2%	88.0%	90.2%	95.7%	96.5%
Overpayment cause (percentage of dollars overpaid)					
Work search	43.6%	51.7%	67.5%	63.2%	45.9%
Benefit year earnings	24.8%	31.3%	14.5%	23.0%	34.0%
Separation issues	13.3%	10.6%	7.6%	10.5%	16.2%
Able and available	2.2%	0.0%	2.0%	0.0%	0.0%
Dependents' allowances	0.0%	0.0%	0.0%	0.0%	0.0%
Base period wage issue	0.3%	0.9%	0.0%	0.4%	0.0%
Severance/vacation/SSI[a]/pension	14.1%	2.4%	4.1%	1.3%	3.4%
Other issues	0.0%	0.7%	1.2%	0.0%	0.0%
Employment services registration	0.0%	0.0%	0.0%	0.0%	0.0%
Other eligibility	1.8%	2.5%	3.2%	1.7%	0.5%

Source: Data are from Improper Payment Information Act (IPIA) one-year data files, available at https://www.dol.gov/general/maps/data.

Note: Annual IPIA figures are based on BAM cases from the 12-month period covering the third quarter of the prior calendar year through the second quarter of the reference calendar year. Overall overpayment rate percentages are from the Improper Payment Rate tab for each year. Overpayment responsibility percentages are from the Overpayment by Responsibility tab for each year. Overpayment cause percentages are from the Overpayment by Cause tab for each year.

[a] Social Security Insurance.

Work search requirements. Claimants in North Carolina are required to contact three employers during the week for which they claim benefits and document their work search contacts. Before July 1, 2018, the number of required work search contacts was slightly higher, with claimants required to make five contacts.

Claims filing options. Most claimants in North Carolina file claims either online or over the phone. Exhibit A.2 shows the proportions of initial and continued claims filed using various methods.

Exhibit A.2. UI claims filing methods in 2018, North Carolina

Type	Sample size	Internet	Telephone	In person	Mail	Employer	Other	Missing
Initial claims	508	72.5%	27.3%	0.0%	0.0%	0.0%	0.2%	0.0%
Continued claims	533	76.7%	22.6%	0.0%	0.2%	0.0%	0.5%	0.0%

Source: Data are from reports generated from https://oui.doleta.gov/unemploy/filingmethods/filingclaims.asp, selecting 2018 as the calendar year, North Carolina as the state, and choosing initial and continued claims in turn. Accessed November 26, 2019.

Note: The output generated by the online report includes this disclaimer: "This report is based on Benefit Accuracy Measurement data for UI claimants with compensated weeks beginning in the period from Q1 2018 to Q4 2018. These percentages are estimated from sample data and are subject to sampling and non-sampling error."

UI administration. In North Carolina, DES administers and verifies UI benefits in the state, and contracts with DWS for reemployment services. After the first claim is paid, each claimant is scheduled for a review within four weeks. For claimants eligible for the RESEA program, the review is conducted during an RESEA appointment. For other claimants, the appointment is conducted via an EAI. The EAI is a shorter meeting, focusing on a review of the claimant's work search log, confirmation of photo identification, and notification of available services. The RESEA meeting is longer and sometimes takes place in a group setting. In either meeting, if DWS staff detect any anomalies in claimant work search logs, they alert DES adjudication staff. Although the state's UI claimant handbook references this in-person verification, it is not featured on the DES website.

B.   Study development and intervention design

When developing the study in North Carolina, our team followed a six-step process for designing and implementing behavioral interventions (Darling et al. 2017), summarized in Exhibit A.3.

Exhibit A.3. Six-step process

As described in Section III.B, we conducted an initial behavioral audit using publicly available information, followed by refining our understanding and developing more specific intervention options in collaboration with state officials during an early site visit to the state in May 2019.

Our work identified five behavioral bottlenecks that could have been contributing to work search errors:

- **Inattention and lack of timely communication.** DES staff believed that claimants did not read the information about work search requirements carefully when filing online. Although the details of the number of required work search contacts are explained simply and clearly on the DES website, the information was nested in an FAQ section that claimants may or may not view. Program and frontline staff indicated their impressions that people navigate through the site quickly.

- **Misunderstanding of work search.** DES staff perceive that claimants do not always understand what counts as a valid work search contact—especially for job searches conducted online. They also perceive that claimants fail to understand that work search must occur every week and/or may not understand how many work search contacts are required. Compliance seemed to improve after the in-person RESEA/EAI meeting, during which workforce staff might point out shortfalls in documentation and refer the case for adjudication.

- **Complexity of maintaining adequate supporting documentation.** DES and DWS staff shared that claimants often do not keep a record of their correspondence with employers, and there is limited guidance on what documentation to keep. Moreover, reporting this information correctly to DES involves additional complexity: Claimants often do not use the form provided for logging work search activities or may use it incorrectly.

- **Lack of understanding of work search review process.** We noted that North Carolina's communications did not emphasize that everyone's work search record would be reviewed in person at either an RESEA or EAI meeting. If claimants were unaware of this requirement, they might doubt whether a verification process exists because North Carolina does not require claimants to submit logs or supporting documentation during the weekly claim certification process. Also, given the short average duration of receiving benefits in North Carolina, the verification steps at the RESEA/EAI meetings typically occur for most claimants only late in their claim period, so claimants may consider the repercussions of poor compliance to be low. If claimants do not believe their work search activities will be checked, the emphasis on keeping records for five years might make some of them overlook the verification that will occur much sooner at the RESEA/EAI meetings.

During our site visit, we also explored the performance and capabilities of current DES systems and procedures, identifying a number of options for the intervention. Our review revealed multiple levers for deploying behaviorally informed communications. They included alerts within the SCUBI claims filing system, emails and letters generated for new claimants, mass point-in-time emails/mailings, and landing pages that appear after log-in and before entering the benefit claims site. Our findings suggested that DES could have the system capabilities to develop an intervention at low cost and to scale successful ones across client touchpoints within the system. When designing the intervention, we considered these key capabilities.

Our team worked with the state over a series of meetings to develop the intervention materials. During these meetings, we discussed how different features could tackle the behavioral barriers we had identified and used a mockup as a springboard for discussion and design. This strategy proved successful in garnering DES support for fielding the intervention, identifying the steps and partnership strategy necessary to launch the study within a month, and iterating it quickly to finalize the design of the materials and get them approved. In subsequent weeks, we worked with DES staff to translate our prototypes into three intervention components: (1) a programmed system alert (Exhibit A.4), (2) an email reminder (Exhibit A.5), and (3) a tip sheet (Exhibit A.6).

Exhibit A.4. Intervention alert

Starts with an upbeat message emphasizing benefits

Makes number of required job contacts prominent

Claimants are required to acknowledge the alert before continuing

Clearly and concisely explains key aspects of work search requirements —uses visual cues

Confirm Alert

Confirm	Alert Message
☐	If you are filing a weekly certification during a week you performed work, you are required to report your gross earnings when you submit your weekly certification, even though you may be paid later. Individuals who knowingly make false statements, or knowingly fail to disclose material facts are committing fraud. UI fraud and abuse are punishable by law. Violators will face penalty and consequence. **Remember to stay on top of your job search!** Starting early and looking for work every week can really pay off-people who apply for work regularly are more likely to find a job quickly. And you need to make **3 job contacts** each week to keep getting UI. Your records of work-search efforts will be reviewed, so make sure to follow these three steps:
☐	🔍 Search. Plan and complete **3 job contacts every week**. 📋 Record. Write down key details about your weekly work search using this form. 📁 Save. Save supporting info (such as emails sent) to provide proof when asked. If your records are incomplete, you may have to pay back your UI benefits. Use these tips to get it right.

Acknowledge

Exhibit A.5. Intervention email reminder

Makes work search requirements salient by communicating about them separately

Starts with an upbeat message emphasizing benefits

Makes number of required job contacts prominent

Reduces barriers to action by providing direct links to tools

Highlights possible consequences of failure to meet requirements

Amin, Samia

From: DoNotReplyDEV@nccommerce.com
Sent: Monday, July 15, 2019 10:04 AM
To: Amin, Samia
Subject: Stay on top of your job search while claiming UI benefits

Dear job seeker,

Starting early and looking for work every week can really pay off—people who apply for work regularly are more likely to find a job quickly. And you need to make **3 job contacts** each week to keep getting UI.

Your records of work-search efforts will be reviewed, so make sure to follow these 3 steps:

- **Search.** Plan and complete **3 job contacts every week**.
- **Record.** Write down key details about your weekly work search using this form.
- **Save.** Save supporting info (such as emails sent) to provide proof when asked.

If your records are incomplete, you may have to pay back your UI benefits.

Need help figuring out what to do? Use these tips to get it right.

Sincerely,

Division of Employment Security
Customer Call Center
P.O. Box 27967
Raleigh, NC 27611
Tele: 1-888-737-0259
Fax: 919-250-4315
Hours: 8:00 AM to 5:00 PM Monday-Friday

Exhibit A.6. Intervention tip sheet

Annotations around the tip sheet:

- Clarifies guidance about what counts for work search and what does not
- Provides a checklist for keeping track of tasks
- Makes key information easy to scan by organizing it by work search contact method
- Gives brief step-by-step instructions in simple language

C. Evaluation design and data sources

As described in Section III.D, to answer key research questions about how the intervention was implemented, and its impacts, we developed a two-phase evaluation design comprising four components: (1) an RCT focused on measuring impacts on claimant knowledge of work search requirements, (2) a descriptive analysis of engagement data, (3) a QED pre-post analysis to analyze impacts on work search behavior, and (4) qualitative research to understand stakeholder perspectives.

The RCT design corresponded with the first of the study's two impact research questions, "Did the intervention change claimants' knowledge of work search requirements?" The QED pre-post analysis yielded evidence on the study's second impact question, "Did the intervention reduce claimants' rates of work search errors and overpayments?" The descriptive analysis of engagement data corresponded with the implementation research question, "Did claimants see the alert, email notifications, and tip sheets?" And the qualitative research corresponded with the first three implementation research questions:

1. What did we learn about our behavioral diagnosis and intervention? What were staff and claimant perspectives on behavioral barriers?

2. Was the intervention applied as intended, and how did staff and UI claimants perceive it?

3. How did the context affect the study and our interpretation?

Below, we describe in greater detail the design of each evaluation component and discuss our data sources.

1.   **Design of each study component**

Our quantitative analyses used different outcomes of interest and samples. Exhibit A.7 summarizes the outcomes, samples, and time frames for each of the analyses. The remainder of this section describes the design of each component in detail.

Exhibit A.7. Outcomes, samples, and time frame of each quantitative evaluation component

	Outcome	Sample	Time frame
RCT	Self-attested work search noncompliance	24,416 claimants randomly assigned (12,262 treatment and 12,154 control) with the treatment group receiving the alert and email (with links to the tip sheet)	July–August 2019
Engagement analysis	Clicks on links embedded in intervention alert and email	7,602 claimants randomly assigned to treatment and had alert set	July–August 2019
QED, NC adjudications	Overpayment related to work search noncompliance	All 183,724 claimants filing from November 2018 to February 2020, including claimants filing before RCT (all comparison group), claimants filing during random assignment (both treatment and comparison group), and claimants filing after RCT (all comparison group)	November 2018– February 2020
QED, BAM	Proportion of work search contacts accepted by BAM audit	461 paid claims (with 1,455 associated work search contacts) in QED, NC adjudications analysis that were also sampled by BAM	November 2018– February 2020

Note: BAM = benefit accuracy measurement; NC = North Carolina; QED = quasi-experimental design; RCT = randomized controlled trial.

RCT focused on claimant knowledge. Claimants who initiated a new claim from June 29 to August 16, 2019, were assigned to two groups based on their claimant IDs. Those with an odd-numbered ID were part of the experimental group and received the three elements of the behavioral intervention. North Carolina's system programmed a work search alert to appear when claimants logged into the online system to file their weekly certifications. Claimants in the treatment group needed to acknowledge the alert by checking a confirmation box and clicking an acknowledgment button before they could proceed to subsequent screens that would allow them to file their certifications and ultimately receive benefits. The alerts appeared at the start of each new claims week for which the claimant was eligible to file. In addition, these claimants received an email sent on either Tuesday or Wednesday of each eligible week with the same information. Those with even-numbered IDs did not receive a programmed alert and emails.

The RCT needed to be short because North Carolina was eager to disseminate the intervention alert and email to all claimants.[11] The brief duration meant that the RCT only randomized 69,761 individuals and could not accumulate a sufficient sample size to detect impacts on the outcome of greatest interest: work search overpayments, as measured either by North Carolina's adjudication system or the BAM program. Instead, the RCT focused only on claimants' self-reports of work search compliance as entered when filing their online claims.

Descriptive analysis of engagement. To help us understand the degree to which claimants in the RCT were exposed to the intervention materials, we planned a descriptive analysis of engagement. The focus of the analysis was to observe the degree of claimant engagement (that is, the number of clicks to links to the online tip sheet in the alert and email).

QED focused on work search behavior. The timing of the intervention's rollout also produced a natural experiment that can be used to assess the impact on outcomes of receiving the work search alert. Exhibit A.8 shows the timeline of the rollout. The state first rolled it out in July 2019 to the treatment group only. It was paused in August 2019, during which time no one received the intervention. The state then rolled it out to all UI applicants in October 2019. We took advantage of this variation in treatment status over time to assess the impact on outcomes of receiving the intervention. This assessment required the assumption that changes over the study period would not have had an impact on the rate of overpayments. Given the on-and-off nature of treatment, and that work search behavior is likely to be more related to an individual understanding of requirements than market conditions, we believe this assumption to be reasonable.

---

[11] During the RCT, each alert was set to have a two-week duration for treatment group members. The number of two-week periods in which the alert was set was a function of how long the claimants' UI spell lasted, among other factors discussed later. Overall, 44 percent of treatment group claimants had the alert set one time, 18 percent two times, 14 percent three times, and 24 percent four or more times. The online system was set such that an email would be sent to claimants upon an alert being set for them.

Exhibit A.8. Timing of work search alert intervention rollout

| | | | Received treatment? | | |
| | | | Claimants with an odd-numbered ID | Claimants with an even-numbered ID | |
Period	Start	End			Analyses drawing on this period's data
Pre-intervention	Nov 2018	Jun 2019	No	No	QED NC adjudications; QED BAM
RCT	Jul 2019	Aug 2019	Yes	No	Self-attested work search noncompliance; engagement
Post-RCT	Aug 2019	Oct 2019	No	No	QED NC adjudications; QED BAM; qualitative analysis on stakeholder perspectives
Broad rollout	Oct 2019	Feb 2020	Yes	Yes	QED NC adjudications; QED BAM; qualitative analysis on stakeholder perspectives

Qualitative analysis of stakeholder perspectives. To understand the perspectives of both state staff and claimants on aspects of work search, the intervention materials, and (for the state) the process of collaborating on the study, we planned a qualitative analysis based on interviews with state staff and focus groups with claimants.

2. Data sources

When designing the evaluation, we focused on using existing data to estimate the impacts of the intervention, in part because time and resource constraints precluded collecting new quantitative data. The data sources we used were as follows:

1. **SCUBI system data on claimant's self-attested work search activities:** When claimants file their weekly certification through North Carolina's online portal, they are asked whether they have conducted qualified work search activities. If they attest that they have not, they are reminded that doing so is a requirement for receiving benefits and are then allowed to change their answer. If they do so, they can still receive benefits for that week. If they do not change their response, an eligibility issue is raised for staff investigation that may result in a denial of benefits.

   Data on claimants' attestations about their work search activities are captured by North Carolina's system. The outcome of interest for the RCT was whether the claimant initially asserted that they did not conduct work search, regardless of whether they subsequently changed their answer when prompted. These data are available for all claimants, and around 8 percent of claimants in the pre-intervention data attested to not conducting work search in at least one of their first eight weeks of claims. For the purpose of examining impacts of the intervention, we excluded the first week the claimant filed: Because a claimant is not exposed to the alert until the first time the claimant files a weekly claim, it is not possible for the alert to have affected the work search behavior (or lack thereof) in the previous week, which is the subject of the first weekly claim.

   We also used data from the SCUBI system for the engagement analysis. The system allowed the state to document each time a user accessed the tip sheet online (that is, web page hits) through a server

log. Moreover, the log allowed us to classify the origin of the hit—whether it came from the alert link or email link.

2. **State adjudications data on improper payments:** Approximately one month after claimants have received their first UI benefit payment, they are asked to attend an RESEA or EAI meeting with DWS staff. At that meeting, claimants must show documentation of their work search activities. If DWS staff believe that work search requirements were not met for a week in which the claimant received benefits, they send a notice to an adjudicator describing their concern. The adjudicator then decides whether to issue a determination for failure to conduct an active work search. However, neither the review by RESEA/EAI staff nor the adjudicator involves verifying work search activities with employers or others.

   The RESEA and EAI meetings cover all claimants with paid claims, rather than a sample of them (like BAM audits). However, an analysis of pre-intervention data revealed that adjudicators identify approximately 1 percent of claimants as ever having received an improper payment due to work search. This low observed rate has two important implications. First, North Carolina's adjudication process identifies a much lower rate of work search improper payments than the BAM program, suggesting that the more thorough BAM review identifies many more improper payments.

   Second, it means that an experimental study would need quite a large sample to detect reasonably sized impacts of the intervention on the rate of improper payment findings by state adjudicators. For example, an impact of only 0.5 percentage points would imply that the intervention changed the outcome of half of the claimants who would otherwise have received an improper payment. The results of other behavioral intervention studies suggest the potential for more modest impacts— approximately 0.15 percentage points (Chojnacki et al. 2017; Darling et al. 2017; Amin et al. 2017); this would require a sample of over 100,000 claimants. Unfortunately, the sample size for the short-term RCT was too small to reliably detect impacts of that magnitude. On the other hand, the sample size for the Phase 2 medium-term QED was larger and does offer sufficient statistical power to estimate impacts on improper payments as determined by state adjudicators.

3. **BAM microdata on acceptable work search contacts:** The official measure of work search improper payments comes from the BAM federal audit program, through which states audit random samples of their claims. For paid claims, the BAM audit involves a close examination of the work search contacts reported by the claimant and includes attempts to follow up with employers to verify the information. The number of BAM audits (roughly 40 per month) is sufficiently small that it was not feasible to analyze claim-level outcomes using BAM data. However, for paid claims, BAM records include fields that identify both the total number of work search contacts made by the claimant and the number of contacts judged acceptable, with roughly four contacts reported per claimant, on average. Because the number of work search contacts in the BAM microdata is roughly four times the number of audited claims, focusing our analysis on the acceptability of individual work search contacts provided us with a large enough sample size for analysis. Using BAM microdata covering paid claims from November 2018 to February 2020, we expanded the data so the unit of observation was an individual work search contact and then constructed an outcome measure equal to 1, if the work search contact was judged acceptable, and zero otherwise. We also linked the BAM microdata to claims data, which allowed us to observe both the timing of the initial claim associated with each BAM paid claim and relevant claimant characteristics (for example, demographics, base period wages, prior employment).

4. **Interviews and focus groups:** Our qualitative analyses are based on data from interviews with state staff and focus groups with claimants. To understand state perspectives on the intervention materials

and their experience in collaborating with our team to design the intervention and implement the evaluation, we interviewed a senior manager and the director of benefit claims at DES in fall 2019, soon after the RCT had been fielded. To understand claimant perspectives on work search requirements and get their feedback on the intervention materials, we convened 10 focus groups with a convenience sample of 20 claimants in February 2020. By this time, the intervention had been scaled to all UI claimants, and our focus group sample should have received the intervention. Focus group participants were claimants who visited the Capital Area NCWorks Career Center in Raleigh over a two-day period in early February 2020 to attend RESEA appointments. During the visit, we also conducted a focus group with six NCWorks frontline case managers and staff.

Statistical power. When designing the RCT and the QED pre-post analysis, we assessed the likely statistical power of each evaluation component to confirm up front that the data would give us sufficient precision to detect impacts on the outcomes of interest. We assessed statistical power by computing minimum detectable impacts (MDIs) for the three outcomes of interest for the RCT and QED analyses. For a given outcome measure, the MDI represents the smallest true impact that the analysis is likely to detect. Relative to the control group means, our pre-study calculations showed MDIs of 5 percent for the self-attestation outcome that was the focus of the RCT, 16 percent for the overpayments outcome that was part of the QED analysis, and 16 percent for the work search contact acceptability outcome that was also part of the QED analysis. In other words, based on our expected sample sizes, we anticipated that the RCT, for example, would have sufficient power to detect the impact of the intervention, if the rates of self-attestation among the treatment group were at least 5 percent greater than the rates among the control group. These MDIs were smaller than the impacts found in other behavioral intervention studies (Chojnacki et al. 2017; Darling et al. 2017; Amin et al. 2017). One note about the MDIs for the QED pre-post analysis is that our calculations assumed a slightly larger sample than we were ultimately able to use; because of the COVID-19 pandemic, we ended the sample window roughly one month earlier than originally planned.

D. Analysis methods

For each evaluation component, we applied methods tailored to the goals of the analysis.

- **RCT analysis.** To confirm that assigning claimants to the treatment or control conditions based on whether their claimant ID was odd or even was a valid random assignment mechanism, we tested for differences between the treatment and control groups, based on their background characteristics. Exhibit A.9 illustrates that only one significant difference was detected out of the 22 hypothesis tests, supporting the assumption that the assignment was random.[12]

---

[12] Under the null hypothesis that the assignment was truly random, the probability of finding at least one statistically significant difference—at the 0.05 significance level and when conducting 22 independent hypothesis tests—is 68 percent.

Exhibit A.9. Baseline equivalence

Measure	Treatment	Control	Difference
Female	52.05%	53.80%	-1.76** pp.
Hispanic	5.59%	5.34%	0.26 pp.
English is preferred language	99.02%	99.19%	-0.17 pp.
Age	42.08	41.85	0.22
Veteran	7.47%	6.96%	0.51 pp.
Disabled	2.80%	2.99%	-0.19 pp.
High school degree, some college	66.61%	67.29%	-0.68 pp.
Bachelor's degree	16.66%	16.97%	-0.31 pp.
Master's or doctorate degree	6.79%	6.38%	0.41 pp.
Base wage	$33,223	$32,962	$260
Number of previous jobs	1.86	1.86	0.01
Average duration of previous jobs (years)	3.16	3.17	-0.01
Any previous full-time job	92.66%	92.28%	0.38 pp.
Missing: Female	0.08%	0.10%	-0.02 pp.
Missing: Hispanic	5.37%	5.17%	0.21 pp.
Missing: English is preferred language	0.12%	0.19%	-0.07 pp.
Missing: Veteran	3.33%	3.31%	0.02 pp.
Missing: Disabled	8.64%	8.49%	0.15 pp.
Missing: Education	0.01%	0.02%	-0.02 pp.
Missing: Base wage	0.73%	0.77%	-0.04 pp.
Missing: Number of previous jobs and average duration	15.19%	14.79%	0.40 pp.
Missing: Any previous full-time job	15.52%	15.12%	0.40 pp.

* Indicates statistically significant at the 0.05 level.

** Indicates statistically significant at the 0.01 level.

In this study, 20 percent of the treatment group—that is, claimants with odd-numbered IDs—did not have the work search alert set and did not receive the intervention email. According to DES, this issue arose because of the timing between when the claimants' benefit week became available and when the software program was run to set the alert (claimants were identified for the email based on whether the alert was set for them). Therefore, we used an instrumental variable (IV) technique to estimate the impact for those claimants who had the alert set and were sent the email. The IV approach is commonly used for situations involving noncompliance, when not all participants randomized to a certain treatment (or control) group actually receive that treatment (see Angrist et al. 1996). In essence, it takes the impact estimated by comparing the treatment and control groups and scales it up to account for noncompliance. Crucially, the IV approach does not assume that the compliers—in this case, claimants with odd-numbered IDs who received the alert and email—are equivalent at baseline to the noncompliers—claimants with odd-numbered IDs who did not receive the alert. Exhibit A.10 displays the rates of compliance for each group.

Exhibit A.10. Receipt of intervention alert and email

Measure	Control	Treatment
Total number of claimants	12,154	12,262
Percentage who had alert/email set	0%	80%
Percentage who acknowledged alert	0%	62%
Percentage who acknowledged alert, of those with alert set	—	78%
Percentage who ever filed a weekly certification	73%	73%
Percentage who ever received a benefit	47%	47%

To increase the precision of the estimates, we controlled for a number of background characteristics with a linear IV regression: age, previous earnings and work history, gender, race, veteran and disability status, and whether English is their preferred language. As a sensitivity test, we also examined impacts when not controlling for background characteristics. We also looked at impacts on alternative constructions of the outcome measure: whether the claimant ever self-reported that they did not conduct work search (that is, including their first certification week) and the number of weeks in which they self-reported that they did not conduct work search. As exploratory analyses, we examined impacts for subgroups, where subgroups are defined by education, previous earnings, and age. We considered these impacts as exploratory, not confirmatory, as defined and recommended by Schochet (2009).

Analysis of engagement. To understand user engagement with the online tip sheet during the RCT, we used the server log data to observe the origin of download requests, or "hits," for the online document. The data allowed us to divide hits into two groups:

1. **Hits from the email link.** One way treatment group claimants could access the tip sheet was to click on the link embedded in the email message sent by DES. For hits from the email link, the server log data identified unique users, who were assigned six-character codes. However, the codes could not be linked to the unique claimant identifier in the claims data, so we could not attribute clicks on the email link to specific claimants.

2. **Hits from the alert link.** The second way claimants in the treatment group could access the tip sheet was by clicking on the link embedded in the SCUBI system alert. For these hits, the server log data contained no information about the specific claimant who had clicked the link.

Using the server log data, we prepared descriptive tabulations of three aspects of user engagement with the tip sheet: (1) the number of hits on the tip sheet, (2) the frequency of multiple clicks on the email link, and (3) the timing of tip sheet access.

QED analysis. For the QED pre-post analysis, we used two different approaches, tailored to two outcomes of interest and their data sources, to estimate the impacts of the intervention materials on (1) the likelihood of an overpayment and (2) the likelihood that an individual work search contact was judged acceptable. Below, we describe separately the methods we used for the analysis of each outcome.

Overpayments analysis

We estimated impact of the intervention on the likelihood of having an overpayment in a given week, for those without a previous overpayment. This allows us to use the full set of data, despite having varying

numbers of weeks of data for each claim. We do this using a hazard model framework, which is designed to account for censored data (Kay 1977). Our hazard model considers the outcome of interest to be the first overpayment observed. Because of when the intervention was in use (Exhibit A.8), we are able to observe individuals in the treatment and comparison groups for different amounts of time. If we simply looked at the presence of overpayments, we would be biased toward finding more overpayments for those groups in the data for longer. In contrast, because the weekly rates of overpayments decrease over the period of a claim, looking at the likelihood of overpayment by week would bias results toward finding more overpayments for groups in the data for less time. The hazard model allows us to avoid introducing either source of bias or having to throw away substantial portions of the data. This analysis implicitly tests for the impact of receiving any overpayment and the time until an individual receives one.

To analyze the hazard model, we use an inverse probability-weighted, stacked logit regression. The stacked logit design describes a hazard analysis that is estimated as a logistic regression on an unbalanced panel data set with an observation for each week in a claim, up to and including the first overpayment. We used this model because it has superior properties for hazard models, based on a small number of discrete time periods. We estimate the following model:

$$Y_{it} = \alpha + \beta T_{it} + \gamma X_{it} + \delta W_t + \varepsilon_{it} \tag{1}$$

Where Y is equal to one if there is an overpayment for individual i in week t, where t is measured relative to the claim file date. T_{it} is an indicator equal to 1 if the individual was treated in week t or in any week before t, X_{it} is a set of individual and week-specific controls, and W_t is a vector equal to 1 in position t and zero in all other positions. β is the coefficient on treatment, which represents the impact of receiving treatment on the conditional likelihood of having an overpayment determined. For ease of interpretation, we translated results into the difference in regression-adjusted average rates of overpayments within five weeks of initial filing.

One additional component of the hazard analysis is that an overpayment in week t is not determined in week t, but instead can be determined for 20 or more weeks following that week. If the data are censoring before the overpayment determination, we may not observe an overpayment that would have been determined over a longer observation period. This situation is problematic because we observed members of the treatment and comparison groups for different amounts of time. To address this issue, we estimated a function, $F(t,x)$, as the probability that an overpayment in week t is determined within x weeks, conditional on the overpayment being determined within 20 weeks. For each x from 0 to 19, we estimated $F(t,x)$ as a regression of an indicator for whether an overpayment was determined within x weeks on t. We estimated the likelihood function for this model as the product of $F(t,x)$ and the likelihood function for the logistic regression model described above. We estimated this model using a maximum likelihood approach. We additionally ran a sensitivity analysis with an adjusted logistic regression and defined the outcome, Y_{it}, as having an overpayment in week t that was determined within 10 weeks. For this specification, we limited the data to claims with at least 10 weeks of data available.

Our sample covered the period from September 2018 to February 2020. Although we received data for March 2020, we excluded these data from the analysis, given the impact of COVID-19 on the unemployment landscape. We limited the analysis to February 2020 so that, for each week of data, we had at least two weeks to observe an overpayment before the start of the COVID-19 pandemic in North Carolina. We limited the analysis to the first claim for each individual in our data. We included a range of

individual characteristics as controls in our analysis, including gender, race, ethnicity, age at initial claim file, education, veteran status, disability status, base wage, preferred language, preferred contact method, number of prior jobs, duration of prior jobs, and whether the individual had previously held a full-time job. Exhibit A.11 shows summary statistics for the treatment and comparison groups, after adjusting for sample weights.

Exhibit A.11. Baseline characteristics of overpayment sample

	Treatment	Comparison	Difference	*p*-value
Gender				
Female	51%	51%	0.1%	0.824
Male	49%	49%	0.0%	0.929
Other	0%	0%	0.0%	0.033
Race				
Black	54%	54%	0.3%	0.480
White	39%	38%	0.2%	0.605
Other	7%	7%	-0.6%	0.016
Hispanic	5%	5%	0.0%	0.974
Age at claim file				
11–30	25%	25%	0.0%	0.942
31–50	48%	48%	0.0%	0.948
51–70	27%	27%	0.0%	0.885
Education				
No high school degree	11%	11%	0.0%	0.964
High school degree	47%	47%	0.0%	0.985
Some college, associate's degree, or certificate	22%	22%	0.0%	0.939
Bachelor's or postgraduate degree	21%	21%	0.0%	0.984
Veteran status	8%	8%	0.0%	0.909
Disabled status	3%	3%	0.0%	0.926
Base wage	$32,702	$32,698	$4	0.977
Preferred language is English	99%	99%	-0.1%	0.076
Preferred contact method: Email	77%	77%	0.0%	0.834
Number of prior jobs				
1	51%	51%	0.1%	0.831
2–4	44%	44%	0.0%	0.931
5 or more	5%	5%	0.0%	0.765
Average duration of prior jobs				
0–2 years	71%	71%	0.0%	0.966
3–9 years	21%	21%	0.0%	0.955
10 or more years	8%	8%	0.0%	0.880
Any prior full-time jobs	93%	93%	-0.7%	0.000

Source: North Carolina claims data.

Work search contact acceptability analysis

One way the intervention could improve work search compliance outcomes would be to lead claimants to successfully complete and document each work search contact well enough to pass a BAM review at a higher rate than claimants who did not receive the alert and email. To estimate the impact of the intervention on the acceptability of work search contacts, we first prepared the raw data for analysis. We received from the state a targeted extract of BAM microdata that covered 700 paid weeks (that is, claims), ranging from the week ending Saturday, October 27, 2018, through the week ending Saturday, February 22, 2020. Our primary outcome of interest was a binary indicator of whether the BAM auditor judged as acceptable an individual work search contact reported by the claimant. From the raw data we received from the state, we dropped 239 paid weeks because they featured missing or zero values for the number of work search contacts, were erroneous duplicates, or could not be matched to a valid new initial claim record.[13] Cleaning the data left us with an analysis sample of 461 paid weeks, each associated with a unique claimant. Expanding the data set by the number of work search contacts reported by each claimant produced an analysis file of 1,455 work search contacts.[14]

We coded each work search contact in the data as being in either the treatment or comparison group based on whether the claimant would have been exposed to the intervention materials. Individual work search contacts were classified as being in the treatment group in the case of either of the following:

1. The BAM paid week fell during or after the week the intervention was rolled out to all claimants (that is, during or after the week ending October 19, 2019).

2. The initial claim associated with the BAM paid week was filed when the RCT was active AND the claimant was assigned to the RCT treatment group.[15]

Exhibit A.12 shows the sample sizes for the treatment and comparison groups for both the number of BAM paid weeks and the number of work search contacts. Because the sample period featured fewer months when the intervention was active, the treatment group is much smaller than the comparison group.

Exhibit A.12. Work search acceptability analysis sample sizes

Group	Number of BAM paid weeks	Number of work search contacts
Treatment	172	547
Comparison	289	908
Total	**461**	**1,455**

Source: North Carolina BAM microdata, payment data, and claims data.

Exhibit A.13 shows a cross-tabulation of the numbers of work search contacts reported by claimants in the sample and the number classified as acceptable.

---

[13] One paid week was missing data on the number of work search contacts, 141 paid weeks reported zero contacts, 5 paid weeks were erroneous duplicates of records from other BAM batch weeks, 82 paid weeks could not be matched to a new initial claim in the extract of claim record data that the state provided, and 10 paid weeks matched only to new initial claims filed after the BAM batch week. Notes in the BAM data indicated that nearly all claimants reporting zero work search contacts were not required to search for work.

[14] As described below, our estimation approach accounts for clustering at the claimant level.

[15] For weeks meeting this condition, we imposed an additional criterion that the initial claim had to have been filed at least two weeks before the end of the RCT period, when the state stopped setting the alert and sending emails.

Exhibit A.13. BAM paid weeks by numbers of work search contacts reported and considered acceptable

Number of work search contacts reported	Number of acceptable contacts						
	Zero	**1**	**2**	**3**	**4**	**5+**	**Total**
1	2	1	0	0	0	0	**3 [< 1%]**
2	2	0	1	0	0	0	**3 [< 1%]**
3	85	72	80	176	0	0	**413 [90%]**
4	6	5	3	2	7	0	**23 [5%]**
5+	6	3	1	4	2	3	**19 [4%]**
Total	**101 [22%]**	**81 [18%]**	**85 [18%]**	**182 [39%]**	**9 [2%]**	**3 [< 1%]**	**461 [100%]**

Source: North Carolina BAM microdata, payment data, and claims data.

Note: Percentages in brackets are the total number of paid claims.

The majority of claimants in the data (90 percent) reported precisely three work search contacts, with few claimants (1 percent) reporting less than three and the remainder reporting more than three. Overall, only 194 claimants in the sample (42 percent) had three or more work search contacts deemed acceptable.[16]

Exhibit A.14 summarizes the characteristics of claimants and paid weeks in our sample.

---

[16] Note that work search contacts not classified as acceptable are classified into two other categories: unacceptable or unverifiable. Only work search contacts classified as unacceptable are considered improper.

Exhibit A.14. Claimant and paid week characteristics, by treatment status

	Treatment	Comparison	Difference†
Female	51.4%	53.9%	-2.7 pp.
African American	15.0%	14.9%	0.1 pp.
Hispanic	3.1%	4.3%	-1.2 pp.
Age			
35 or younger	24.5%	23.9%	-0.6 pp.
36–45	22.9%	25.2%	-2.3 pp.
46–55	25.6%	22.3%	-3.3 pp.
55 or older	27.1%	28.6%	-1.5 pp.
Educational attainment			
No high school degree	10.8%	7.4%	3.4 pp.
High school degree	38.8%	38.4%	0.4 pp.
Some college, associate's degree, or certificate	24.0%	26.5%	-2.5 pp.
Bachelor's degree or postgraduate degree	26.5%	27.6%	-1.1 pp.
Veteran	9.0%	9.1%	-0.1 pp.
Disabled	2.2%	2.6%	-0.4 pp.
Preferred contact method is email	75.5%	73.9%	1.6 pp.
Base period wages	$38,434	$39,389	-$955
Number of previous jobs	1.8	1.8	0.0
Duration since claim file date (weeks) ***			
1–4	30.2%	29.6%	0.6 pp.
5–8	32.2%	29.2%	3.0 pp.
9–11	14.8%	16.4%	-1.6 pp.
12–13	14.1%	8.9%	5.2 pp.
14 or more	8.8%	15.9%	-7.1 pp.
Missing values:			
African American	69.3%	70.9%	-1.6 pp.
Hispanic	5.7%	4.0%	-1.7 pp.
Veteran	3.5%	4.6%	-1.1 pp.
Disabled	6.6%	11.0%	-4.4 pp.***
Number of observations	**547**	**908**	

Source: North Carolina BAM microdata, payment data, and claims data.

Note: ***/**/* Statistically significant at the 1/5/10 percent level. †Tests of statistical significance are based on t-tests for continuous measures and χ^2 tests for binary and categorical measures. For categorical measures, statistical significance is noted for the overall measure, where applicable.

Even before any potential adjustments, the treatment and comparison groups were fairly well balanced. The differences between the two groups were statistically significant for only two characteristics—the duration since the initial claim file date and missing information on disability status. Among the set of duration measures, the imbalance was driven in large part by differences between the two groups in the proportions of audited claims representing weeks at least three months after the initial claim filing date.

Exhibit A.15 summarizes the outcome data for the treatment and comparison groups.

Exhibit A.15. Proportion of work search contacts classified as acceptable, by treatment status

Work search contact disposition	Treatment	Comparison	Difference†
Acceptable			
Yes	61.8%	57.0%	4.8 pp.*
No	38.2%	43.1%	-4.9 pp.*

Source: North Carolina BAM microdata, payment data, and claims data.

Note: ***/**/* Statistically significant at the 1/5/10 percent level. †Test of statistical significance is based on a χ^2 test; differences are unequal due to rounding.

The data show that the proportion of work search contacts judged acceptable was slightly higher among the claims in the treatment group—nearly 62 percent of the work search contacts reported by these claimants were considered acceptable, compared to 57 percent among those reported by comparison group claimants. The difference is marginally statistically significant (at the 10 percent level), though the difference does not necessarily reflect the causal impact of the intervention.

To estimate the impact of the intervention on the likelihood that a single work search contact was judged acceptable, our main specification was to estimate a regression model using an inverse probability weighting (IPW) procedure. Compared to a simple linear regression approach, implementing IPW seeks to eliminate the potential for bias in the estimated treatment effect stemming from compositional differences between the two groups (Mansournia and Altman 2016). Estimation was a two-step process. In the first step, we estimated a logistic regression with the treatment indicator as the dependent variable, as in equation (2):

$$Treat_{ic} = \alpha_1 + \beta_{10} X_i + \beta_{11} Z_c + \varepsilon_{1ic} \tag{2}$$

In the equation, i indexes individual claimants, c indexes work search contacts for a given individual, $Treat_{ic}$ is an indicator equal to 1 if the claimant was exposed to the intervention and zero otherwise, X_i are controls for the claimant characteristics listed in Exhibit A.14, Z_c are the duration controls listed in Exhibit A.14, and ε_{1ic} is an error term. After estimating the model, we used the estimated coefficients to produce a treatment propensity score for each observation.

The second step was to estimate the average treatment effect using a weighted linear regression of the outcome on a constant and a treatment indicator, with the weight for each observation equal to the inverse of its propensity score from the first step. The regression equation is given by equation (3):

$$Y_{ic}^A = \alpha_2 + \delta_{1PW} Treat_{ic} + \varepsilon_{2ic} \tag{3}$$

In the equation, Y_{ic}^A is an indicator equal to 1 if the work search contact was acceptable, and zero otherwise. After the first step, we assessed the balance between the treatment and comparison groups by estimating a series of weighted regression models, such as equation (3), but for which the dependent variable was replaced sequentially with each control variable. If the estimated difference between the two groups (that is, δ_{1PW}) was statistically significant, the result would suggest that the weights were

inadequate to balance the groups for the covariate in question. Such a result could suggest adding the covariate as a control variable in equation (3) to control for the remaining imbalance. However, none of the estimated differences was statistically significant, even at the 10 percent level. This confirms that the weights adequately balanced the treatment and comparison groups for every individual covariate; therefore, when estimating impacts, we did not add control variables to equation (3). Finally, recognizing the likely correlation between the acceptability of work search contacts made by the same claimant, we accounted for clustering at the claimant level.

Implementation analysis. To glean insights from the interviews and focus groups, we transcribed our notes and reviewed them to identify common themes. The study team transcribed and analyzed the focus group discussion data using a coding scheme to identify themes and areas of variation. The coding system included categories related to awareness of and feedback on intervention materials, knowledge of work search requirements, perceptions of ease of adherence to work search processes, determinants of work search, and suggestions for process improvements. Because of the small number of interviews and focus group participants, the goal of our analysis was not to produce generalizable findings. Rather, we sought to use the qualitative research to supplement and enhance the quantitative findings and help us consider the most promising avenues for future studies to build off our work.

Notes on interpreting the results. As noted in Section III.E.3, there are at least two important considerations to bear in mind when interpreting our findings. First, the perspectives of claimants from our focus groups may not be generalizable. Second, our QED analysis may not isolate the impacts of the intervention from the impacts of other concurrent factors affecting the outcomes.

E. Supplemental findings

1. Implementation analysis

Analysis of engagement. Our analysis of the engagement data focused on understanding how claimants in the treatment group during the RCT engaged with the online tip sheet. Exhibit A.16 summarizes the number of claimants in the treatment group who viewed the tip sheet web page, by origin (that is, via the system alert or email link).

Exhibit A.16. Tip sheet hits, by origin

Source	Number of tip sheet hits	As percentage of treatment group
Email link (unique users)	681	7%
Alert link (total, unique users unknown)	1,459	15%
Total (assuming all hits are unique)	**2,140**	**22%**

Source: DES tip sheet server log data.

Note: Percentages are calculated by dividing the number of tip sheet hits in the server log data by the number of claimants in the treatment group who received the alert and email (9,771).

Only 681 unique users, representing just 7 percent of the treatment group, clicked on the link in the email. The alert generated more clicks than the email, but the data do not allow us to identify whether these are

unique visitors, nor the degree to which they overlap with those who accessed the link via email.[17] If all clicks on the alerts were from distinct users, they would amount to 15 percent of users. Overall, the data suggest that no more than 22 percent of claimants in the treatment group viewed the tip sheet.

Our analysis of the timing of claimants accessing the tip sheet suggests that those who engaged did so within the first few days of the alert being set and the emails delivered. Exhibit A.17 shows the frequency of hits on the tip sheet that originated from the email link; Exhibit A.18 shows the frequency of hits from the alert. Most of the hits occurred on Mondays and Tuesdays, when the emails were sent. The daily volume of hits from the email link decreased rapidly within a couple of days. Hits from the alert link showed similar patterns of decay, with most hits occurring from Saturday to Monday.

Exhibit A.17. Number of clicks on the email link to the tip sheet, by day of the week
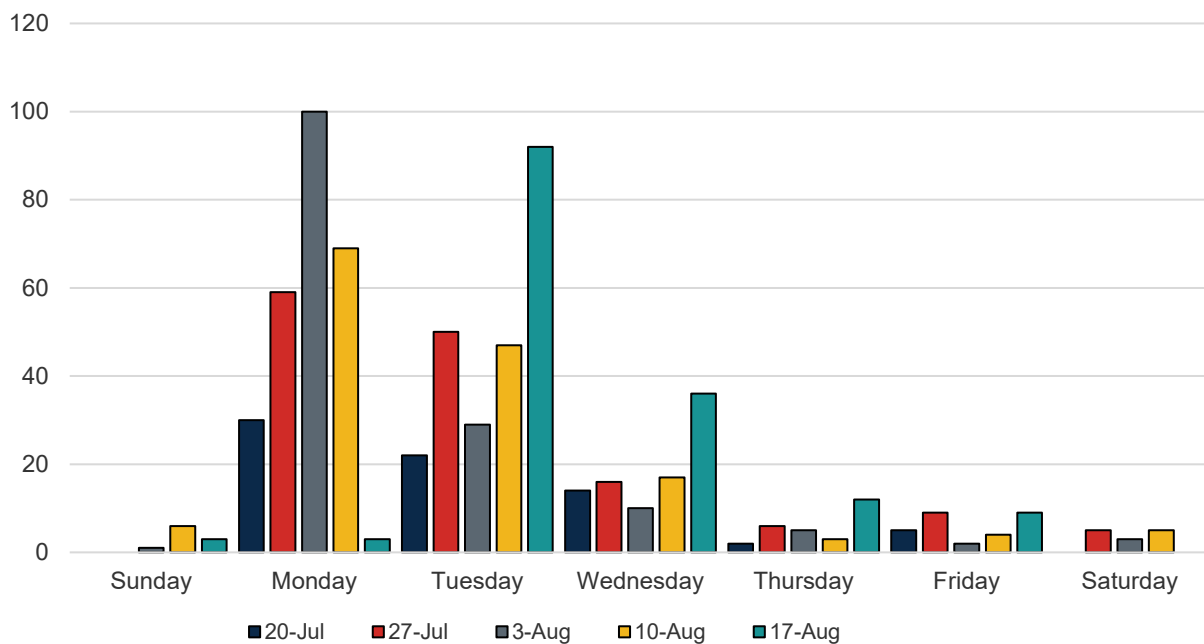
Source: DES tip sheet server log data.

Note: Data are grouped by the Saturday that ended each week, as identified in the chart legend. Emails were delivered by DES on Monday or Tuesday each week.

---

[17] Because the weekly alert no longer appears after the claimant acknowledges it, such repeat clicks would have had to occur in different weeks.

Exhibit A.18. Number of clicks on the alert link to the tip sheet, by day of the week
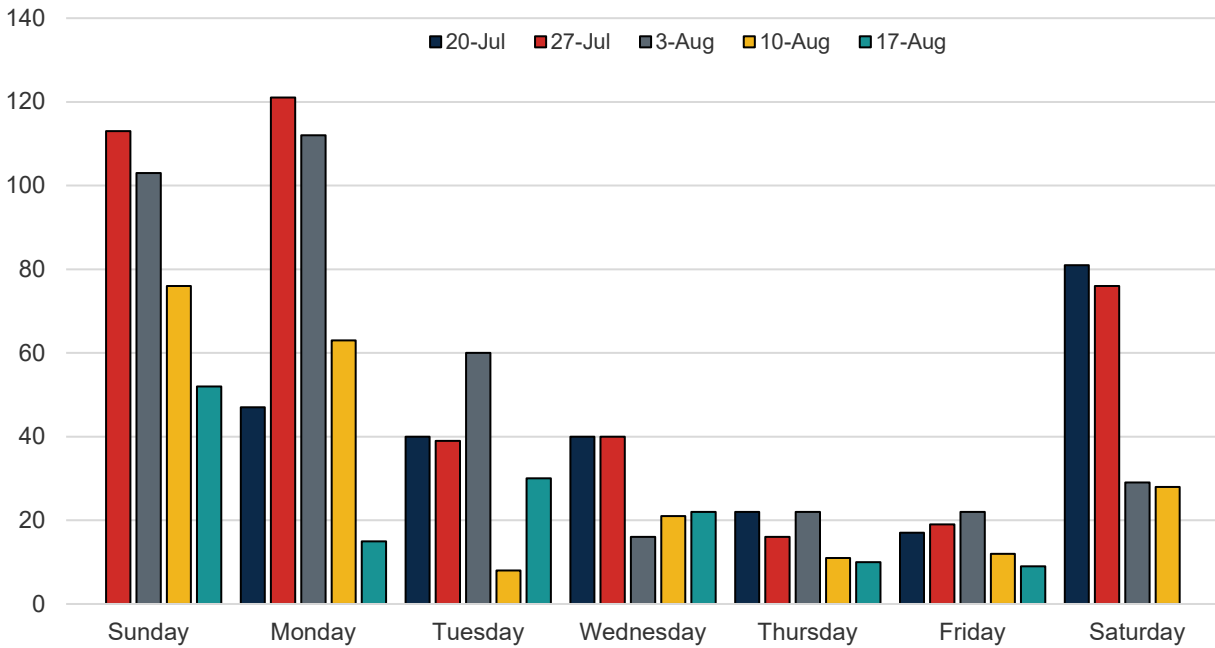
Source: DES tip sheet server log data.

Note: Data are grouped by the Saturday that ended each week, as identified in the chart legend. Alerts were set when a new benefit week became available to the claimant to file against, at 12:01 a.m. each Sunday.

We also assessed the degree to which claimants who accessed the tip sheet found it useful enough to revisit. Exhibit A.19 shows the distribution of the number of times the same user visited the tip sheet using the email link.

Exhibit A.19. Distribution of the number of clicks on the email link, per unique user
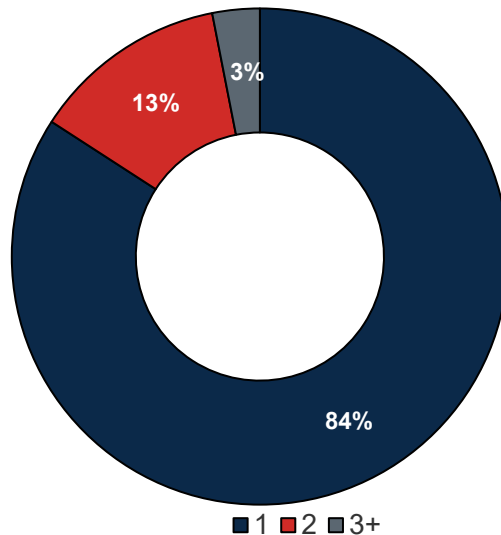
Source: DES tip sheet server log data.

Note: We calculated the number of unique users, excluding multiple clicks by the same user in a single day, which we interpreted as a single engagement episode. Separate calculations show that 45 percent of repeat clicks within a single day occurred within 1 minute of each other, and only 31 percent were separated by at least 30 minutes.

The data show that only a small fraction of users clicking the email link revisited the tip sheet by clicking the link again. Only 13 percent of the 681 unique users who clicked the email link did so twice, and 3 percent clicked more than twice. In total, only 16 percent of these unique users who clicked on the email link did so multiple times in the same email. Because we were not able to identify specific claimants in the data, our counts of repeated clicks omitted, for example, claimants who clicked on email links in two different weekly emails.

Perspectives of state staff. We spoke with staff in North Carolina to gather their feedback on the intervention and understand their views of the experience of collaborating to implement the project. Through our interviews, we learned the following:

1. **DES leadership found the intervention materials responsive to their perception of the causes of work search errors.** In their view, claimants misunderstanding when to begin work search, how to document it, and how long to retain their records were the main drivers of work search errors.

2. **State leadership thought the intervention materials were well designed.** The two DES staff members with whom we spoke had positive views on the intervention materials. Specific features cited as appealing included (1) using concise and clear language; (2) integrating icons and visuals to appeal to visual learners; and (3) using design effectively, especially on the second page of the tip sheet.

3. **The state scaled the use of the intervention materials.** The state used the intervention materials in a broader effort to reduce work search errors. Three ways it used the materials were to (1) incorporate posters displaying the materials into training sessions for DWS staff, (2) program the alert and send reminder emails to all claimants, and (3) display large posters of the tip sheet at American Job Centers.

4. **The state was interested in exploring the application of behavioral insights to other interactions with claimants, both through the UI office and DWS.** DES staff reached out to the study team to brainstorm additional ways they could apply insights from behavioral science to improve their operations. They were especially interested in making behavioral insights-informed revisions to their work search log and integrating materials from the intervention into invitations for the RESEA program. They were also open to working with DWS to see whether behavioral interventions could be used to improve the RESEA program experience.

5. **The state found that collaborating to design and implement the intervention was low cost, but collaborating on the project was more time intensive than they had anticipated.** Although DES leadership did not explicitly track the costs associated with implementing the project, they felt the effort required for implementation was minimal. However, DES staff noted that partnering on the evaluation component of the project required more effort than originally envisioned, citing, in particular, the work around data sharing and the need to collaborate over a longer-term period (12 months in lieu of 6). DES also had anticipated receiving the impact findings more quickly.

Claimant perspectives. We spoke with 20 claimants in focus group discussions about their perspectives on work search requirements and their thoughts on the intervention materials. Through the focus groups, we learned the following:

1. **Claimants understood the number of required work search contacts.** Those with whom we spoke indicated they were making many more than three work search contacts per week and did not need a reminder about the required number. That number seemed reasonable to them.

2. **Claimants were less clear about documentation and record retention requirements.** Claimants used a range of documentation methods, including the form provided by the state or their own electronic or paper records. Yet others indicated they had created the work search log expressly for the RESEA meeting. Several claimants indicated they were applying for jobs through platforms like Indeed or LinkedIn, which allowed them to track their activity and/or had online/email folders with their application summaries. Some of them knew they needed to keep records for five years.

3. **Claimants felt the log templates provided by the state were cumbersome and did not aid their work search efforts.** We heard that creating a work search log was not helpful to planning and executing the search, and that maintaining the log was a distraction from it. Respondents indicated they were actively and consistently looking for work. A few respondents said that maintaining a log felt like busywork that detracted from their job search processes. They also indicated that making the required number of work search contacts was not hard. Rather, what they found difficult was continuing to apply for positions without hearing back from employers, going through multiple rounds of interactions with employers without success, and staying optimistic while worrying about how to make ends meet.

4. **Some claimants noted the effects of communication tone on morale and motivation for work search.** Claimants explained that it was difficult to maintain an optimistic outlook and continue work search, despite being unsuccessful in finding a job. Some claimants indicated that the sometimes-harsh tone of communications from the state exacerbated the challenge.

5. **Some claimants referenced structural elements that affected work search.** Some claimants told us that the low level of benefits (North Carolina's maximum benefit amount is $350) gave them a sense of urgency regarding work search.

6. **Claimants recalled seeing the alert, but said they only skimmed it.** Virtually all of the claimants with whom we spoke said they saw the system alert when they logged in, but most appeared not to have read the information in it very closely. None had noticed or clicked on the embedded link to the tip sheet. Respondents explained that they were in a rush to acknowledge the alert so they could move on to filing their claims. Some respondents indicated the font was small in the pop-up alert. A few claimants indicated that they had filed their claim using the web browser on their phones, which could have made the alert especially hard to read.

7. **Few claimants recalled seeing the email.** Several claimants said they received a lot of email communications from the UI office and DWS—including survey invitations—so they tended not to pay much attention to messages from those senders.

8. **Virtually no claimants had been exposed to the tip sheet, though they felt it looked useful.** At most, only a couple of the claimants with whom we spoke may have seen the tip sheet. When given a copy to review, claimants preferred its format and content over other instructions they had received. Specific feedback from claimants on the tip sheet included the following:

   o The content was useful, and more intuitively and attractively presented than the instructions they had received.

   o The incorporation of color was inviting and contrasted with the low-resolution work search log.

   o Suggestions for improving the content of the tip sheet included adding more explicit guidance on how to log work search contacts made through job application portals, such as Indeed. One individual suggested removing language indicating that a job search would not be valid if it was for a position for which the claimant did not have knowledge, skills, and abilities, because it could deter people from considering jobs in other industries for which they had relevant skills.

9. **Claimants shared that exposure to the tip sheet, as early and often as possible, would be useful.** After reviewing the tip sheet, several claimants indicated that it was useful. Some of the claimants with whom we spoke suggested that receiving a copy of it early, such as during the waiting week, would have been useful. There were mixed opinions regarding the best way to share the tip sheet—whether by mail, email, both email and mail, or embedding in the SCUBI system.

2. Impact analysis

RCT analysis

The estimated difference between the treatment and control groups in the likelihood of claimants self-attesting that they did not complete work search was 0.2 percentage points—not statistically significant. Exhibit A.20 displays the impacts, in which the treatment group mean represents the control group mean plus the estimated impact based on the IV regression, which accounts for noncompliance and controls for baseline characteristics.

Exhibit A.20. Impacts on self-reports of work search noncompliance
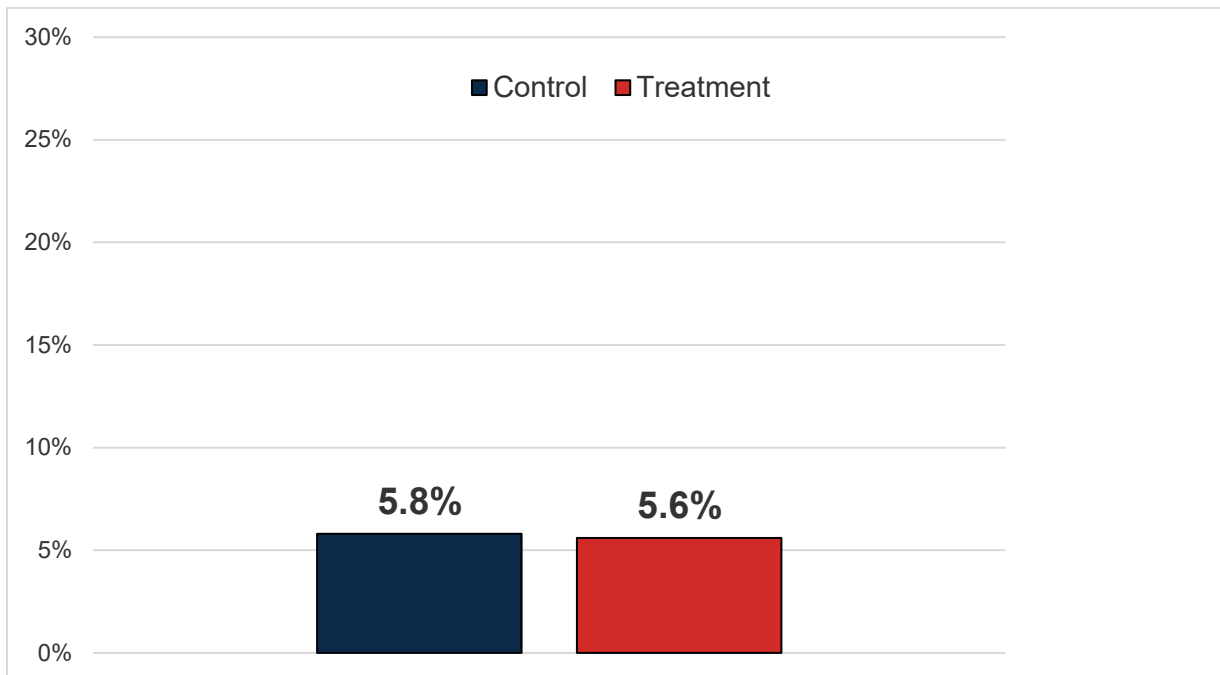
Note: The control group mean represents the unadjusted proportion of claimants who ever reported they did not conduct work search, excluding the first claimed benefit week. The treatment group mean represents the control group mean, plus the estimated impact from the IV regression. The difference between groups was not statistically significant.

We also found no impacts when we examined two alternative versions of the outcome measure: (1) ever self-reporting a failure to conduct work search (*including* the first claimed week) and (2) the number of weeks the claimant self-reported that they did not conduct work search (Exhibit A.21). Finally, there were no impacts for any subgroups in which we categorized claimants by education, age, and prior earnings. Again, this outcome measure captures one specific feature of claimant knowledge regarding work search requirements: whether they self-report that they did not conduct work search, even though this would lead to their weekly benefit claim being denied. However, it does not necessarily measure other aspects of claimants' knowledge or beliefs about work search requirements or their actual work search behavior.

Exhibit A.21. Impacts on alternative outcome measures

Measure	Treatment	Control	Difference	Standard error
Ever self-reported no work search, excluding first claimed benefit week (main outcome)	5.62%	5.80%	-0.18 pp.	0.37
Ever self-reported no work search	8.22%	7.92%	0.31 pp.	0.44
Number of self-reports of no work search	0.10	0.09	<0.00	0.01

* Indicates statistically significant at the 0.05 level.
** Indicates statistically significant at the 0.01 level.

As shown in Exhibit A.22, there were no statistically significant impacts on any of the subgroups defined by education, age, and prior earnings. However, there was little statistical power available to detect subgroup impacts.

Exhibit A.22. Subgroup impacts

Measure for all groups	Group	Treatment	Control	Difference	Standard error	*p*-value
Ever self-reported no work search, excluding first claimed benefit week (main outcome)	Does not hold a bachelor's degree	5.72%	5.67%	0.05 pp.	0.43	0.909
	Bachelor's or other advanced degree	5.30%	6.24%	-0.94 pp.	0.73	0.197
	Under 40 years of age	5.37%	5.26%	0.11 pp.	0.51	0.830
	40 years of age or older	5.84%	6.30%	-0.46 pp.	0.54	0.397
	Below median in prior earnings	5.16%	4.51%	0.65 pp.	0.50	0.197
	Above median in prior earnings	6.13%	7.11%	-0.98 pp.	0.55	0.076

QED analysis. As noted in Section III.E.2, our QED pre-post analyses found no impacts on either overpayments, as measured by state adjudication data, or the likelihood that a work search contact was deemed acceptable, based on BAM microdata.

Overpayments analysis

- We found no impact of the intervention on overpayments. Exhibit A.23 shows the regression-adjusted likelihood of having an overpayment determined for one of the first five weeks of filing the initial UI claim. We chose to illustrate this time period for tractability, and because 72 percent of overpayments are determined within five weeks. On average, 0.54 percent of individuals in the comparison group had an overpayment determination relative to 0.51 percent of the treated group. This difference was not statistically significant and does not represent a meaningful decrease in overpayments from a policy perspective.

Exhibit A.23. Regression-adjusted likelihood of having an overpayment in the first five weeks of UI (20-week observation period)
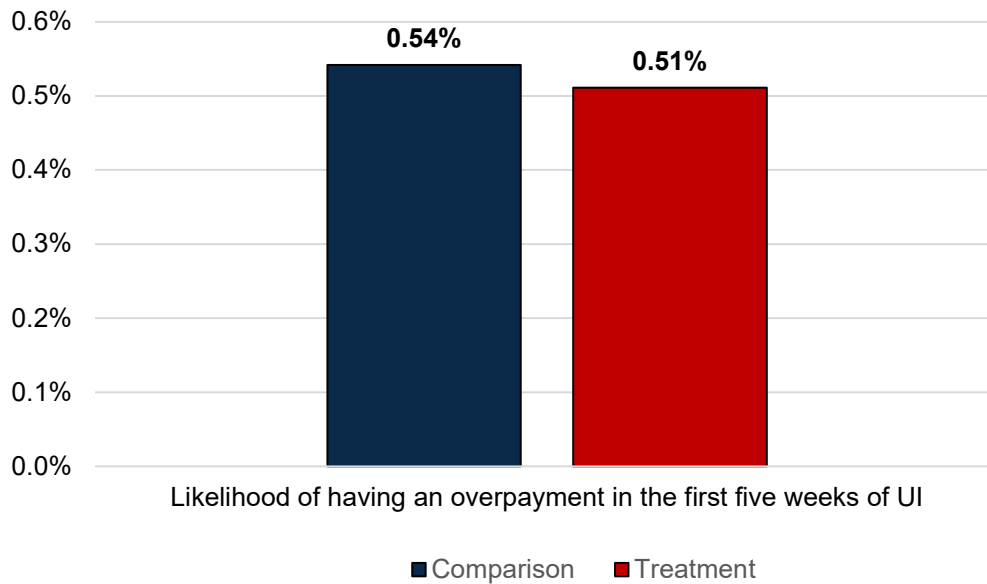
Likelihood of having an overpayment in the first five weeks of UI

■ Comparison   ■ Treatment

Note: The difference between the treatment and comparison groups was not statistically different, *p* = .44. Results represent the regression-adjusted impact of having an overpayment within five weeks of the UI claim file date. Results are limited to overpayments determined within 20 weeks of occurrence. All weeks that occurred during the treatment period, or for which a previous week within the same claim had been in the treatment period, were considered treated. The difference between groups was not statistically significant.

- As a sensitivity analysis, we restricted the analysis to overpayments determined within 10 weeks. Exhibit A.24 shows the results of this analysis as regression-adjusted averages. As expected, the rates of overpayments are lower in this sensitivity because we restricted our observation window. However, this sensitivity analysis is consistent in showing no statistically significant or meaningful difference in the average rates of overpayments between the treatment and control groups.

Exhibit A.24. Regression-adjusted likelihood of having an overpayment in the first five weeks of UI (10-week observation period)
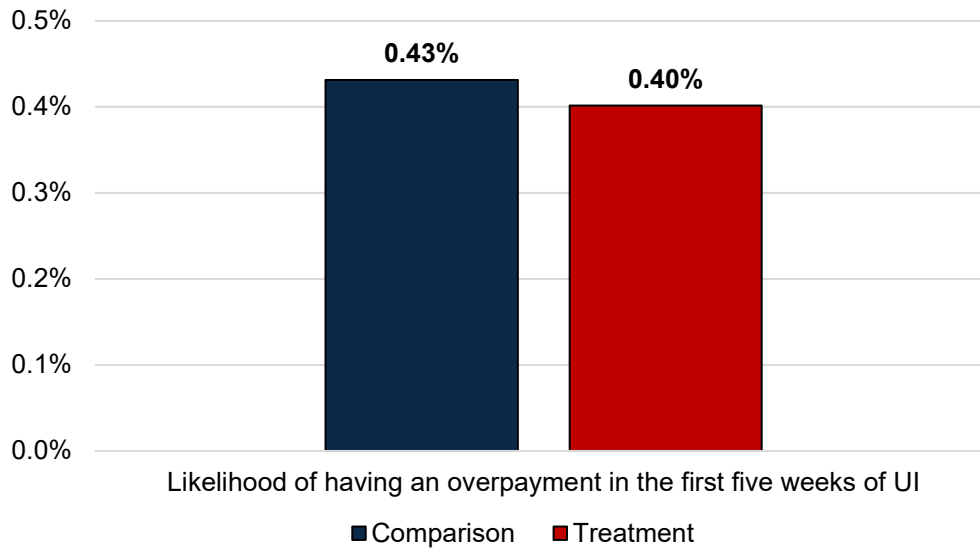
Note: The difference between the treatment and comparison groups was not statistically different, *p* = .45. Results represent the regression-adjusted impact of having an overpayment within five weeks of the UI claim file date. Results are limited to overpayments determined within 10 weeks of occurrence. All weeks that occurred during the treatment period, or for which a previous week within the same claim had been in the treatment period, were considered treated. The difference between groups was not statistically significant.

It is important to note the overpayment rate of about 0.43 percent in the comparison group—the rate of work search related overpayments, as detected by North Carolina's adjudication system—is substantially lower than the rate of improper payments due to work search as detected by the BAM program (Table A.1). Although the BAM does not report a rate defined in precisely the same way, it finds that improper payment rates in North Carolina represent 16 to 24 percent of paid claims, depending on the year, and that one-half to two-thirds of the improperly paid dollars are due to work search issues.

Work search contact acceptability analysis

Exhibit A.25 shows the estimated regression coefficients from the IPW regression model (equations [2] and [3]) we used to estimate the impact of the intervention on the likelihood that an individual work search contact was acceptable.

Exhibit A.25. IPW regression estimates, work search contact acceptability

Estimating equation	Variable	Coefficient	Standard error
(2)	Female	-0.056	0.214
	African American	-0.123	0.388
	Hispanic	-0.303	0.601
	Age		
	35 or younger[a]	-	-
	36–45	-0.154	0.293
	46–55	0.078	0.295
	55 or older	-0.041	0.300
	Educational attainment		
	No high school degree	0.464	0.381
	High school degree[a]	-	-
	Some college, associate's degree, or certificate	-0.145	0.269
	Bachelor's degree or postgraduate degree	-0.021	0.285
	Veteran	0.184	0.365
	Disabled	-0.117	0.685
	Preferred contact method is email	0.201	0.264
	Base period wages (thousands)	-0.002	0.004
	Number of previous jobs	-0.017	0.093
	Duration since claim file date (weeks)		
	1–4 [a]	-	-
	5–8	0.107	0.263
	9–11	-0.090	0.318
	12–13	0.447	0.356
	14 or more	-0.581	0.358
	Missing values		
	African American	-0.142	0.297
	Hispanic	0.728	0.487
	Veteran	0.230	0.632
	Disabled	-0.810*	0.470
	Constant	-0.340	0.511
(3)	Treatment	**0.064***	**0.035**
	Constant	0.569	0.024

Source: North Carolina BAM microdata, payment data, and claims data.

Note: ***/**/* Statistically significant at the 1/5/10 percent level. Coefficients for equation (2) are logit coefficients that represent changes in the predicted log odds that are associated with each one-unit increase in the corresponding explanatory variable, holding all other explanatory variables constant; coefficients for equation (3) are based on ordinary least squares regression. The sample included 1,455 observations of work search contacts made by 461 claimants.

[a] Reference category; omitted from the regression.

Our main finding for this outcome is the estimated coefficient on the treatment indicator in equation (4), which shows that the intervention is estimated to have increased the likelihood that a single work search contact was acceptable by 6.4 percentage points—a result that is statistically significant only at the 10 percent level. (The fact that only one of the 22 coefficients in equation [2] is statistically significant supports the idea that the observed background characteristics did not influence whether claimants ended up in the treatment group rather than the comparison group.)

As robustness checks, we also estimated impacts on the work search acceptability outcome using both a linear probability model and a logistic regression. Equation (4) shows the specification for the linear probability model.

$$Y_{ic}^A = \alpha_3 + \beta_{30} X_i + \delta_{LPM} \text{Treat}_{ic} + \varepsilon_{3ic} \tag{4}$$

The logistic regression fits a logit model using the same control variables as those in equation (4). For the linear probability model, the parameter of interest is δ_{LPM}, which represents the marginal effect of being exposed to the intervention on the likelihood that the work search contact was acceptable. For the logit model, the coefficient on the treatment indicator is an odds ratio, which we convert into an estimated marginal effect comparable to the results from the other specifications. Exhibit A.26 summarizes our impact estimates for the linear probability model and logit regression specifications, along with the IPW estimate for comparison. Our estimates are consistent across the three specifications, which produce virtually identical results.

Exhibit A.26. Alternative specifications, work search contact acceptability

	IPW regression model	Linear probability model	Logit regression model
Impact	0.064* (0.035)	0.061* (0.035)	0.061* (0.035)
Control variables			
Claimant demographics	No	Yes	Yes
Claimant employment history	No	Yes	Yes
Claim duration	No	Yes	Yes
Missing value indicators	No	Yes	Yes
Number of observations	1,455	1,455	1,455
Number of claimants	461	461	461

Source: North Carolina BAM microdata, payment data, and claims data.

Note: ***/**/* Statistically significant at the 1/5/10 percent level. Standard errors are reported in parentheses. The IPW regression model did not include control variables because the weights sufficiently balanced the treatment and comparison groups.

Finally, to explore whether the impact on work search acceptability varied with selected claimant characteristics, we estimated a modified version of equation (1) that incorporated interaction terms between the treatment indicator and indicators for three subgroups of interest: (1) educational attainment of a bachelor's degree or higher, (2) age 40 or older, and (3) base period wages greater than or equal to the sample median ($32,316). Because of the complexity in estimating IPW models that have subgroup interactions, we used a linear probability model to estimate differential impacts for each subgroup of

interest. Each model was like equation (4) but featured an indicator for the subgroup of interest (G_i), along with an interaction between the subgroup and treatment indicators, as shown in equation (5).

$$Y_{ic}^A = \alpha_4 + \beta_{40} X_i + \beta_{41} Z_c + \beta_{42} G_i + \delta_{41} \text{Treat}_{ic} + \delta_{42} \left(G_i \times \text{Treat}_{ic} \right) + \varepsilon_{4ic} \qquad (5)$$

For the first two subgroups, the corresponding categorical variables in the regression (that is, educational attainment categories and age groups) were replaced with the subgroup indicator variable. Exhibit A.27 summarizes the results.

Exhibit A.27. Exploratory subgroup impact estimates, work search contact acceptability

	Bachelor's degree or higher	Age 40 or older	Base period wages greater than or equal to median†
Impact			
In subgroup	0.149**	0.053	0.120***
	(0.063)	(0.043)	(0.042)
Not in subgroup	0.019	0.071	-0.051
	(0.043)	(0.061)	(0.062)
Control variables			
Claimant demographics	Yes	Yes	Yes
Claimant employment history	Yes	Yes	Yes
Claim duration	Yes	Yes	Yes
Missing value indicators	Yes	Yes	Yes
Number of observations	1,455	1,455	1,455
Number of claimants	461	461	461

Source: North Carolina BAM microdata, payment data, and claims data.

Note: ***/**/* Statistically significant at the 1/5/10 percent level. Standard errors reported in parentheses. Each estimate was produced by estimating the linear probability model given in equation (5).

†The difference in impact estimates between claimants in and not in the subgroup is statistically significant at the 5 percent level.

Exhibit A.28. Exploratory subgroup impact estimates, overpayment analysis

	Bachelor's degree or higher†	Age 40 or older	Base period wages greater than or equal to median†
Impact			
In subgroup	-0.449*	-0.043	-0.124
	(0.197)	(0.098)	(0.096)
Not in subgroup	0.011	-0.082	0.026
	(0.083)	(0.114)	(0.117)
Covariates included	Yes	Yes	Yes
Number of observations	3,410,627	3,410,627	3,410,627
Number of claimants	183,724	183,724	183,724

Source: North Carolina claims data.

Note: ***/**/* Statistically significant at the 1/5/10 percent level. Standard errors reported in parentheses. Each estimate was produced by estimating the hazard model given in equation (2), adjusted for covariate interactions as in equation (5). Results are limited to overpayments determined within 20 weeks of occurrence. All weeks that occurred during the treatment period, or for which a previous week within the same claim had been in the treatment period, were considered treated.

†The difference in impact estimates between claimants in and not in the subgroup is statistically significant at the 5 percent level.

The exploratory subgroup analysis shows that the impact of the intervention on the acceptability of work search contacts was especially high for two subgroups. For claimants with a bachelor's degree or higher, the intervention increased the likelihood that an individual work search contact was acceptable by 15 percentage points; the estimate was statistically significant at the 5 percent level. This result was consistent with the overpayment analysis, which found that the intervention decreased overpayments for individuals with at least a bachelor's degree by 0.15 percentage points, although these results were only significant at the 10 percent level. For claimants whose base period wages were higher than the sample median, the intervention had an impact similar in magnitude, equal to 12 percentage points, and that was statistically significant (at the 1 percent level). This result was not found in the overpayment analysis. We also tested whether the difference in impact estimates for each of the three subgroups considered was statistically significant—it was true only for the groups defined by base period wages. The results support our hypothesis that the impacts could differ by claimants' level of education, which could influence their understanding of work search requirements in the status quo condition, but they do not provide any evidence of greater impacts on younger claimants who we speculated might be more adept at interacting with web-based intervention materials. Lastly, the result that the intervention was more effective for claimants with higher base period wages, a proxy measure of socioeconomic background, sheds more light on the types of claimants for whom the intervention may be particularly effective.

Notes on interpretation

There are two important considerations to bear in mind when interpreting our findings:

1. **Perspectives from focus group participants may not be generalizable.** Our conversations with claimants gave us good insights into their perspectives on work search and our intervention materials. However, because we spoke only with those claimants who attended their scheduled RESEA appointments, what we heard may not represent sentiment among claimants more broadly.

2. **Our QED analysis may not isolate the impacts of the intervention materials alone.** Our QED analysis compared outcomes for claimants exposed to the intervention materials to those not exposed, largely leveraging the timing of the state's broad rollout of the intervention materials to all claimants. Although our statistical approaches controlled for observable factors like claimant characteristics, our impact estimates may reflect other unobserved factors that concurrently affected the outcomes.

APPENDIX B
Washington Technical Appendix

A. State context

Washington prioritizes preventing UI work search errors

Washington places a high priority on dealing with work search issues, which, on a dollar basis, account for the majority of UI overpayments for claimants in the state, as estimated by the federally standardized BAM process (Exhibit B.1). Even before engagement with the DOLBI team, state agency staff had begun considering how program innovations might reduce work search errors. This made them an ideal, eager partner for the effort described here.

Exhibit B.1. Annual selected improper payment outcomes in Washington State and nationwide, 2015–2019

Measure	2015	2016	2017	2018	2019
Overpayment rate					
National	10.3%	11.1%	12.1%	12.5%	10.2%
Washington	16.9%	13.6%	8.7%	19.3%	12.6%
Share of overpayments caused by work search errors (percentage of dollars overpaid)					
National	29.4%	37.5%	36.9%	40.1%	32.9%
Washington	62.3%	60.9%	65.2%	60.6%	69.3%

Source: Improper Payment Information Act (IPIA) one-year data files, available at https://www.dol.gov/general/maps/data.

Note: Annual IPIA figures are based on BAM cases from the 12-month period covering the third quarter of the prior calendar year through the second quarter of the reference calendar year. Overall overpayment rate percentages are from the Improper Payment Rate tab for each year. Overpayment responsibility percentages are from the Overpayment by Responsibility tab for each year. Overpayment cause percentages are from the Overpayment by Cause tab for each year. National figures include the 50 states, plus the District of Columbia and Puerto Rico, except for 2015, when Florida was excluded due to inadequate BAM sampling.

Informing claimants about work search requirements. During normal operations of the UI system, all claimants in Washington receive a brief description of their work search requirements as part of the benefits rights document provided at the point of application. All claimants also have access to a complete description of the requirements in the Unemployed Worker Handbook, which is available online (Employment Security Department 2019).

Claims filing in Washington. Unemployed workers in Washington file their initial claim after being laid off. Even if they are found eligible to receive benefits, the benefit week following their initial claim is deemed a "waiting week," for which they cannot receive benefits. They are, however, required to contact employers during their waiting week. Claims for a benefit week cannot be filed until after that calendar week has ended (that is, until at least Sunday of the following week), though they need not be filed immediately after the benefit week ends. The following table from Washington's Unemployed Worker Handbook, Exhibit B.2, provides an example of the claim process and work search requirements. We considered these features of the Washington claims process when designing the intervention.

Exhibit B.2. Example of Washington claims process

Week	Action	Payment
Week 1	Apply for unemployment benefits and begin your job search Example: Jane was laid off on Monday the 1st and applies for unemployment benefits that day. She starts to look for work and writes her contacts down on her job search log	If otherwise eligible for benefits, week one is a ***waiting week***. No benefits are issued for it.
Week 2	Submit weekly claim for week 1 Example: On Sunday the 7th, Jane submits her first weekly claim for the week that ended on Saturday the 6th	If otherwise eligible for benefits, week one is a ***waiting week***. No benefits are issued for it. Even though no payment is issued for this week, you must still submit the weekly claim and meet all eligibility requirements.
Week 3	Submit weekly claim for week 2 Example: On Sunday the 14th, Jane submits a weekly claim for the week that ended on Saturday the 13th.	If otherwise eligible for benefits, the first payment is issued for week 2. **Note:** payments may be delayed while deciding your eligibility.

Source: Washington State Employment Security Department, Unemployed Worker Handbook, January 2019. Available at https://esdorchardstorage.blob.core.windows.net/esdwa/Default/ESDWAGOV/Unemployment/ESD-Handbook-for-Unemployed-Workers.pdf. Accessed February 12, 2021.

Changes to implementation that occurred after the intervention design was complete

During the project, two circumstances affected the study context and our approach to collecting data and analyzing outcomes.

First, in the weeks before the evaluation launch in December 2019, the ESD team notified the study team that, starting in January 2020, ESD would add system prompts to its online filing platform, requiring claimants to document their work search activities before proceeding.[18] Specifically, the system would prevent people from completing their weekly claim until they either (1) submitted contact information for at least three employers whom they had contacted for work search or (2) attested that they had not completed at least three work search activities. People who submitted their claim by phone were not subject to the same requirement, although this difference was likely not apparent to claimants unless they tried using both systems. After launching this system modification, ESD added functionality that allowed claimants to download their log entries as an aid to complying with record-keeping requirements. Both of these changes aimed to lower the prevalence of work search errors by making it easier (and required) for

---

[18] The study team was aware that this change in requirements was in development, but at the outset of evaluation planning in fall 2019, it appeared likely that the change would be implemented either late in the evaluation period or after the evaluation had ended.

the majority of claimants who file online to document their work search activities. Because they made it more difficult for claimants to file a weekly claim without completing their work search activities, the changes likely would lower the prevalence of work search errors, even in the absence of the email tested in this study. The reporting requirement also made the work search requirement more salient earlier in the claim process than it had been before, thereby addressing one of the key behavioral bottlenecks targeted by the email intervention. As a result, these changes might have diminished the potential scope for the email to further reduce the prevalence of work search errors.

Second, in February 2020, Washington State emerged as one of the earliest and most serious centers of the COVID-19 pandemic in the United States. On February 29, the governor's office declared a statewide emergency, followed by 35 additional proclamations in March alone, related to the virus response. Many of these proclamations encouraged or mandated measures such as business closures and stay-at-home orders. The number of UI initial claims filed in Washington in the four weeks ending April 4 increased by more than 20 times from the previous four weeks (ETA Form 539). On March 24, the governor formally waived the work search requirement, with the change retroactively effective March 8 (ESD 2020). By the time the governor announced the work search waiver, ESD had already responded to the sharp increase in unemployment claims by shifting staff priorities, including suspending JSR activities for claim weeks after March 14 to process claims as rapidly as possible.

More information on the updated data gathering and analysis processes for the study appears below. The pandemic also affected the general context for the evaluation because many claimants likely became preoccupied with personal health and safety concerns, thus affecting their ability and motivation to look for work, whereas those who continued looking for work faced sharply reduced demand for labor in many major sectors.

B. Study development and intervention design

The study and intervention were designed based on a behavioral diagnosis described above in Section IV.B.

The study team wished to design an intervention based on insights from the behavioral science literature that would also be feasible to implement at low cost. A key goal was to create an intervention that provided timely information to increase claimant understanding of work search requirements. During our site visit to Washington, we explored the performance and capabilities of current ESD systems and procedures, identifying multiple options for a behavioral intervention. Several elements of Washington's existing procedures and the online claims and adjudication system permitted us to conduct this study quickly and at low cost. The study team worked with ESD to determine low-cost ways to provide information on work search requirements and found that emails to targeted groups of claimants would be feasible, whereas changes to the claimant online filing portal would not.

In particular, emails can be programmed to be sent to claimants selectively, based on parameters such as the date of filing or the last digit of their Social Security number (our method of assignment to the treatment group). Also, the online system permits claimant-level linkages of data across systems, including the claim file and JSR and other adjudication data.

During the site visit, we worked with ESD to determine the most feasible levers to target that would use a low-cost intervention and afford opportunities for rigorous learning. We learned that any intervention that involved programming messages into the online claim filing portal ("eServices") would not be feasible in

the evaluation's time frame. However, the state's claims data and email campaign system were compatible with an intervention to send emails to a new group of claimants each week.

The intervention email for this study—an example of which is shown in Exhibit IV.2—was sent on the Tuesday morning following a claimant's first filed weekly claim—that is, the claim for the claimants' "waiting week," for which they are unable to receive benefits. Therefore, most claimants received the intervention email early in the first week for which they might be eligible to receive benefits.

As an alternative, we considered sending the email immediately after a claimant submitted the initial claim but decided against it for several reasons. First, to avoid drop-off in the impact over time, we wanted the message to arrive at a time when the claimant was required to search for work to be eligible for monetary payment. In our analysis sample, about 6 percent of claimants received their treatment email three or more weeks after filing their initial claim because they did not always begin filing immediately after that initial claim. Second, a large portion of initial claims never progress to benefit receipt, and we did not want to reduce the measured impact of the intervention by applying it to a subgroup who would never be subject to work search requirements. Third, if we had included all initial claimants in the sample but excluded from our analysis those who never filed waiting week claims, we would have excluded any claimants who responded to the treatment by forgoing weekly claims altogether, when we wished instead to ensure we measured any such response to the treatment.

In place of the high-level summary of requirements and informational hyperlinks contained in the intervention email, the study team also considered attaching or linking to a tip sheet that would provide additional information on the details a claimant must report for each type of job contact. However, the evaluation time frame coincided with a change in policy regarding the precise information required for each job contact, raising the possibility that a tip sheet shared at the beginning of the evaluation period could be misleading for the same claimant later in the period. To avoid this possibility, the study team removed the tip sheet from the intervention design.

C. Evaluation design and data sources

To assess the impacts of the intervention email, we gathered data on several topics, including treatment group members' engagement with the treatment email, key outcomes of interest, and claimant characteristics that could help explain variation in outcomes unrelated to the experimental treatment.

Engagement with the treatment email

The automated system that sent treatment emails recorded recipients' engagement four hours after the emails were sent and—for a subset of study weeks—again seven days later. The system recorded the total number of unique individuals who opened each message and clicked each link, without identifying which specific sample members opened or clicked the emails.

Adjudication data and outcome variables

Our initial evaluation design called for measuring overpayments using work search errors identified through ESD's JSR process. Due to the COVID-19 pandemic, both our sample intake period and the follow-up period for claimants were curtailed, and JSRs were halted in mid-March 2020. Instead, we opted to use metrics generated using data from the state's claimant tracking and adjudication systems. This system records weekly payments made to claimants, along with an indicator for those payments subsequently disqualified. It also tracks any work search issues associated with each weekly claim and

any determination the agency makes about whether these issues can be cleared or declared true errors. Individuals who did not show up in this issue tracking system were assumed not to have had any work search issues or errors on any of their claims.

We included in our sample all benefit weeks ending on or before February 29, 2020, because later weeks had their work search requirements waived, albeit in some cases retroactively. All claimants in the sample had the opportunity to file a claim for at least one benefit week after filing their waiting week claim.

These data sets provided four outcomes of interest, all aggregated to the claimant level as a binary indicator, set to 1 for claimants who experienced the outcome in question after receiving their treatment email, and zero for all other claimants:[19]

1. **Work search errors.** This outcome was our primary one, because these errors are the underlying driver for overpayments. They are sometimes identified during a claim review by a person, but also commonly result from issues in the claim itself, such as when claimants indicate they did not search for work.

2. **Payments made and later disqualified.** These payments are similar to overpayments, but not identical to the BAM overpayments measure. We also cannot specifically identify which payments were disqualified due to work search errors.

3. **Work search issues later cleared.** They can be regarded as a sort of "nuisance" outcome, possibly caused by people doing something wrong when filing their claim. They may require staff time to investigate and clear, and at least some may be preventable through information interventions. Accordingly, we measured these issues to assess whether the additional information provided by the intervention email made these errors less prevalent.

4. **Work search errors identified among claimants selected for JSR review.** This outcome is similar to what would have been our main outcome in the absence of COVID-19, and though it includes system-identified errors, it is more heavily weighted toward errors uncovered by human review. This outcome is available only for the subset of claimants sampled for a JSR review—a group substantially smaller than we initially expected, due to the truncation of our study period related to COVID-19.

Because adjudication involves review, ESD's determinations as to a claim's status can change over time. Our outcomes are based on claim status as of April 15, 2020, balancing the reality of ESD's limited review capacity in March and April against the possibility that, for instance, some observations currently declared to be in error might have subsequently been reassigned as cleared issues in a different environment (for instance, following more a careful hand review, or if claimants made successful appeals of unfavorable determinations) or vice versa (if a hand review turned up issues on those claims not currently flagged as problematic).

To help explore potential channels through which claimants react to the treatment email, we also gathered data on the total amount of benefits paid (which would be affected if treatment group members claimed fewer weeks of benefits on average), the likelihood that claimants switched from Internet to phone filing of claims (because that method involves less immediate scrutiny of their job search records), and the likelihood of applying for work search-exempt "standby" status following the treatment email. To increase the precision of our impact estimates and account for any chance imbalance between the treatment and control groups, we also collected four claimant characteristics: (1) indicators for the county

---

[19] Some paid claim weeks may have preceded the treatment email—for example, if an individual filed late for the waiting week.

associated with the addresses they provided to ESD, (2) an indicator for Spanish language preference, (3) indicators for the claimant's pre-layoff occupation or occupations, based on the three-digit Standard Occupational Classification code assigned using information supplied by the claimant about their prior job, and (4) indicators for the waiting week associated with the claim. We included these groups of variables in our statistical models as regression controls.

For these control variables and several other pre-randomization characteristics, we checked for any evidence suggesting our treatment and control groups were identical before being randomized into the study. For binary characteristics (such as Spanish language preference), we regressed the variables on a treatment indicator. However, for categorical characteristics—study intake week, county geographic location, and occupation—we assessed baseline equivalence using a chi-squared test, rather than testing equivalence on multiple binary indicators. The chi-squared test produces a *p*-value that can be interpreted the same as a *p*-value from a regression—that is, as the likelihood the difference would have resulted by random chance if treatment were assigned at random and the underlying distribution of treatment and control group characteristics were the same. None of the chi-squared tests of our categorical variables indicated a significant difference at the *p* = 0.10 level.

The results of this baseline equivalence analysis are shown in Exhibit B.3. (For study intake date, we show the distribution of the treatment and control groups across the categories without assessing statistical significance for each comparison. The other two categorical variables include too many groups to display.) None of the baseline characteristic differences between control and treatment groups exceeded a tenth of the control group average.

One of the differences (on the likelihood of providing ESD with a valid email address) yielded a *p*-value of less than 0.1. This does not by itself indicate underlying imbalance between the groups because, given the number of characteristics we tested, some would be expected to have low *p*-values by random chance. Additionally, we assessed this difference within the full sample of claimants but screened out of the analysis sample those individuals not providing a valid email address, so this small difference does not threaten the internal validity of the impact estimates. In other words, the findings can be interpreted as representing the impacts on individuals who provide a valid email address, as all members of our analysis sample meet that criterion.

Exhibit B.3. Baseline equivalence of treatment and control groups

	Group means (percent)		Difference (percentage points)
	Treatment	Control	Difference
Provided a valid email address to ESD	89.19	88.58	0.62*
Spanish language preference	3.72	3.80	-0.08
Selected for JSR sample	4.79	4.44	0.35
Sample intake date (Tuesday following waiting week filing)			N/A
Dec 17, 2019	10.35	10.06	0.29
Dec 24, 2019	7.65	7.93	-0.27
Dec 31, 2019	7.94	7.25	0.69
Jan 07, 2020	7.36	7.61	-0.25
Jan 14, 2020	11.47	11.54	-0.07

	Group means (percent)		Difference (percentage points)
Jan 21, 2020	10.57	10.52	0.05
Jan 28, 2020	10.15	10.53	-0.38
Feb 04, 2020	8.72	9.07	-0.36
Feb 11, 2020	9.60	9.44	0.16
Feb 18, 2020	7.91	7.59	0.32
Feb 25, 2020	8.28	8.46	-0.18
Standby application filed before randomization	5.04	4.71	0.33
County of address provided to ESD			N/A
Pre-separation occupation (three-digit SOC)			N/A

Source: ESD claims data.

Note: ***/**/* Statistically significant at the 1/5/10 percent level. Group means shown list the percentage of members of each group to whom the characteristic applies. We estimated differences and statistical significance by regressing the characteristic indicators on a treatment indicator. The sample for the valid email characteristic includes all 30,340 claimants who filed their waiting week claim in time to be added to the sample for the February 25, 2020, entry date, whereas we obtained the sample for other comparisons (26,967 claimants) by eliminating those without a valid email address. For sample intake week, county, and occupation categorical variables, we assessed equivalence and statistical significance using a chi-squared test, the results of which indicated no statistically significant differences between the groups in these characteristics. Individual sample entry weeks are listed, but we excluded categories for county and occupation because of a high number of categories. We did not assess statistical significance for the sample intake week comparisons because baseline equivalence is assessed at the categorical level.

ESD = (Washington State) Employment Security Department; JSR = (Washington State) Job Search Review; SOC = Standard Occupational Classification.

Statistical power. When designing the intervention, we aimed to achieve an MDI of 18 percent for our outcome of interest at the time—errors detected through ESD's job search reviews. This impact is lower in percentage terms than the size of the impacts observed for other recent email interventions focusing on behavioral communications to workers addressing employment- or personal finance-related subject matter (Darling et al. 2017; Amin et al. 2017), and it was expected to be achievable within a reasonable time frame. Based on our expectation about the number of reviews each week, and based on 2018 data suggesting that 22 percent of claimants in the JSR sample would have a work search error, a trial length of 28 weeks was estimated to be sufficient to achieve this MDI. In other words, we anticipated that, if the evaluation had run for its intended duration, it would have had sufficient power to detect the impact of the intervention, if the rate of JSR errors among the treatment group were at least 18 percent—or 4 percentage points—lower or higher than among the control group.

D.   Analysis methods

Implementation data collection and analysis methods

To answer the implementation research questions, the study team conducted semi-structured discussions with five Washington ESD program managers and staff who participated in designing and implementing the email intervention. We conducted these discussions in October 2020. The five staff members were key personnel involved in the diagnosis, design, and implementation of the email in collaboration with the study team. Each discussion lasted approximately one hour. The discussions gathered information on four primary topics aligned to the implementation research questions:

1. **The level of effort and inputs involved in designing and implementing the intervention.** We first asked staff to describe who was involved in designing the intervention, as well as the level of effort associated with developing the intervention. We probed to understand staff perspectives on the key components of the intervention and how ESD implemented them.

2. **The degree to which ESD implemented the intervention as intended.** We next asked staff to share their perspectives on the degree to which ESD implemented the intervention as intended. Staff shared their recollections of whether any changes were made to the email intervention after it was launched, their perspectives on the level of effort involved in implementation, and factors that facilitated or impeded it.

3. **Key features of the context in which the intervention was implemented.** The third topic of focus involved contextual factors that may have affected the evaluation, particularly the email's effectiveness or the evaluation's ability to measure the effect of the email. This section of the conversation explored policy changes that occurred within ESD and staff perspectives on the effects of the COVID-19 pandemic on ESD and the claimant population.

4. **Lessons learned through the design and implementation processes.** Finally, we asked staff to share lessons they had learned or best practices they had identified through the intervention design and implementation processes.

The study team learned additional details about the work search log requirement from program documents and through more discussions with ESD staff, which occurred outside of the semi-structured discussions.

In the analysis, we transcribed data from the semi-structured discussions and coded them to identify themes and areas of variation. After transcribing the data, we used Excel to code those themes present in the interviews. To identify themes, we first pulled relevant data for each question; we then looked across respondents and coded themes for each question. We also identified notable areas of variation, as well as singular responses that appeared important to the context or may have reflected the unique responsibilities of the staff member.

After coding themes by discussion question, we identified overarching themes and insights. We elevated themes that did one or more of the following: (1) provided context helpful to understanding the quantitative findings, particularly the lack of impacts on quantitative outcome measures; (2) shared ESD staff recommendations that would be critical for other state agencies considering similar interventions; and (3) illustrated ESD staff reflections on the ongoing impact on the agency of participating in the intervention. We used these themes to generate potential insights for supplementing and informing the interpretation of the quantitative results.

Analysis of engagement. To understand user engagement with the online tip sheet during the RCT, we measured the number of treatment group members who opened the emails within a predetermined window of time, and the number of recipients who opened each link.

Impact analysis
Our main analysis uses regression analysis to measure the intervention's impact. This method isolates the effects of assignment to the treatment group after separately accounting for the influence of other measured claimant characteristics, such as geographical location and previous job. Because this study was designed as a randomized, controlled experiment and our analysis takes place at the claimant level, the coefficient on the treatment indicator can be interpreted as the average change in each outcome caused by sending the intervention email to a claimant. (Random assignment of treatment status eliminates the

concern about incorrect impact estimates that might result if individuals self-selected into the treatment group.) Our main impact estimates use the ordinary least squares (OLS) model shown in equation 6 below.

$$Y_i = \alpha + \beta_0 X_i + \delta_{LPM} \operatorname{Treat}_i + \varepsilon_i \qquad (6)$$

In equation (6), Y_i represents the outcome measure of interest, X_i represents regression controls included to improve the estimates' precision—specifically, a claimant's county, pre-separation occupation, timing of waiting week, and language preference— δ_{LPM} represents the treatment effect estimate, and ε_i represents a random claimant-level error term.

Most of our outcomes for this study are binary—that is, they are limited to taking on values of either zero or 1. For instance, our main outcome equals 1 for individuals who commit a work search error and zero for claimants who do not (though we rescale our results to run from zero to 100). For binary outcomes, OLS regression is sometimes also called the "linear probability model," and impact estimates can be interpreted as the percentage point change in the likelihood of a given outcome caused by the treatment email.

As a robustness test, we re-analyzed our primary and secondary outcomes using two alternate models—probit and logit. These models sometimes characterize changes to outcomes of very low (or very high) probability more accurately than the linear probability model. Although the numerical magnitude of their coefficients is more complicated to interpret, their direction and *p*-values can be interpreted the same as in OLS models—as an increase or decrease in the likelihood of an outcome, with a *p*-value that reveals whether the change meets statistical significance thresholds. One note to consider when interpreting our probit and logit results is that, because of challenges in incorporating certain low-frequency regression controls in these models, we have excluded from them the nontreatment regression control variables included in our main analysis.

Finally, we also estimated treatment effects for primary and secondary outcomes separately for individuals who filed using different methods. We separated claimants into groups for those who filed exclusively by the Internet, exclusively by phone, and a mix of the two. This approach could help reveal treatment effects that might be obscured if treatment affected the likelihood of errors for some but not all filing methods (such as if, hypothetically, changes to the online filing procedure led online filers to switch to phone filing so they would not have to report their work search contacts). This robustness check used OLS regression with control covariates like the main analysis but allowed the treatment effect to vary by filer type, within a single regression equation. For these models, we report the rate for the base group (Internet-only filers), along with coefficients showing both the average differences between filer groups and the effects of receiving the treatment email on members of each group.

E. Supplemental findings

1. Implementation analysis

We spoke with key Washington ESD staff to learn about their experiences participating in designing and implementing the email intervention. In addition to the key findings described in Chapter IV, we obtained the following supplemental findings through the semi-structured discussions with key Washington ESD staff:

External support increased ESD staff's motivation to adjust the agency's communication approach and test the intervention.

- **ESD staff were motivated to test the intervention because of their interest in improving their UI overpayment rate.** Staff reported that this broad agency goal was a facilitator to implementation.

- **Additionally, funding from DOL and simultaneous engagement from NASWA around behavioral science and UI overpayments were helpful aids in developing and implementing the email.**

Designing the intervention required collaboration and problem solving across the organization.

- **Staff involved in designing the intervention represented multiple Washington ESD teams.** Respondents reported that key staff included the program integrity manager, compliance manager, policy unit, communications team (including the communications manager), UI operations team, and information technology (IT) team. Some staff reported that the communications team approved the final communications and used its resources to disseminate revised information.

- **Washington ESD's IT team had to develop a process for disseminating the intervention.** According to all staff members with whom we spoke, the IT team did not previously have a process for identifying groups of claimants to receive different emails. As a result, the team had to determine how to direct the appropriate emails to treatment and control claimants using existing IT resources.

Members of the IT team were most heavily involved in implementing the email.

- **All respondents confirmed that the IT team's senior data analyst was responsible for implementing the email.** To do so, the staff member used SQL to identify customers as treatment or control claimants, loaded the emails into the GovDelivery system, and collected and shared data on email open and link-clicking rates with the study team. He employed his knowledge of SQL and UI program requirements to successfully implement the email. ESD also designated a second staff member who could send the email, in case the senior data analyst was absent.

- **ESD involved other teams and resources as needed.** ESD also used its data security tools and email subscriber list to implement the email. The JSR team was not involved in sending the email but responded to customer questions that it may have prompted.

The implementation of the intervention occurred as intended, requiring a level of effort similar to what staff anticipated at the outset.

- **Although it required a significant level of effort to determine how to send the email, most staff felt that the level required for implementation was similar to what they had expected.** One area that required more time than anticipated was identifying technical issues related to the functioning of the mass email platform and implementing solutions—though these technical issues only affected a small number of email recipients.

- **Most ESD staff stated there were no changes to the email content or how it was sent once the intervention started.** However, the email provided a hyperlink to a work search log, and even though the hyperlink and email text did not change, some staff reported that the work search log changed during the intervention.

Analysis of engagement. Our analysis of engagement data explored whether treatment group members received the intervention messages and the information they contained. Exhibit B.4 below shows the seven-day engagement figures for treatment group members. Because the system recorded link-opening

statistics separately for each link in the treatment email, we cannot precisely identify the number of individuals who clicked on *any* link.[20] Instead, we present a lower bound, using the most-clicked link. Exhibit B.5 shows the share of recipients who clicked on each of the four most popular links within seven days of receiving the message, noting possible overlap to the extent that some individuals clicked more than one link.
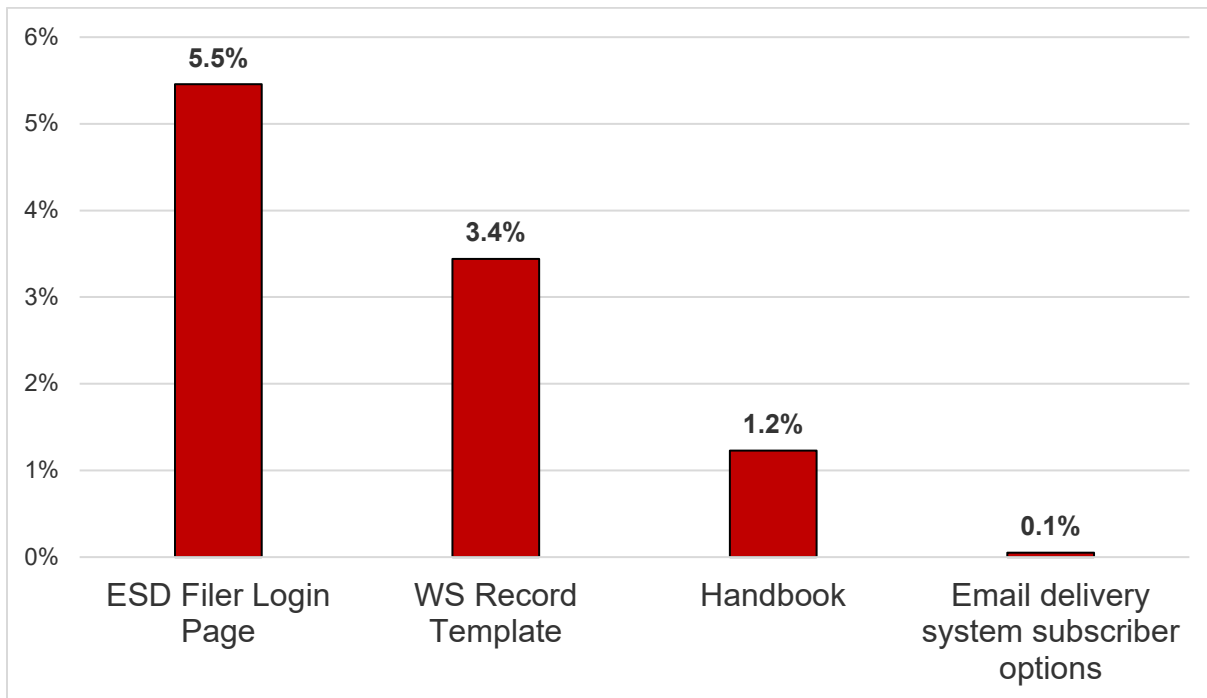
Exhibit B.4. Share of treatment group members who opened their email and those who clicked a link
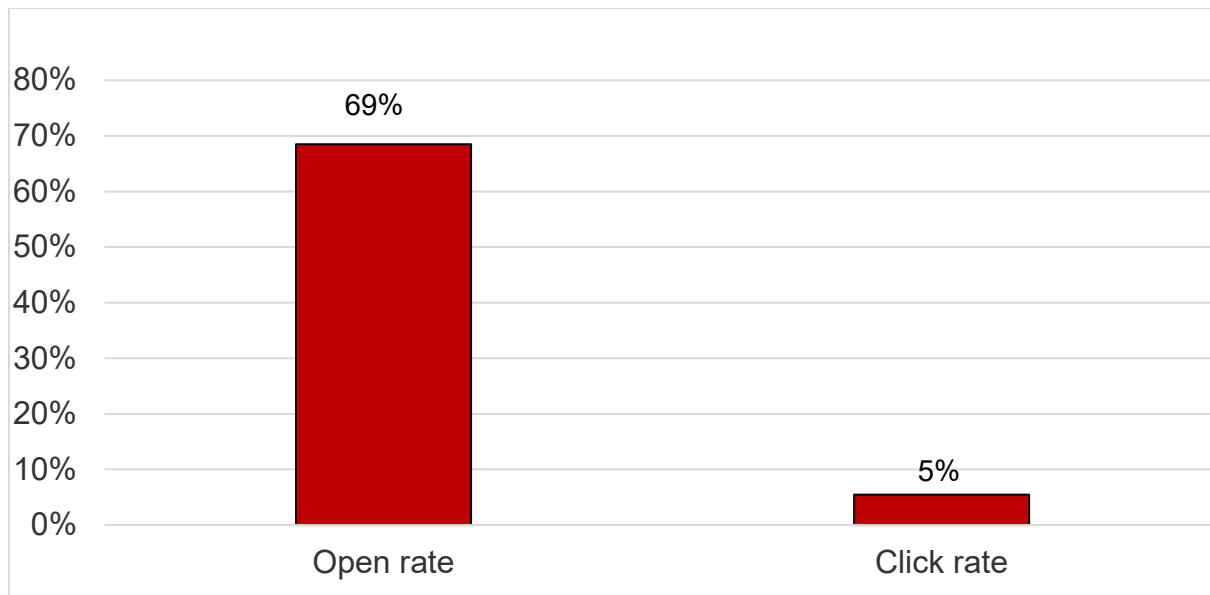
Source: ESD seven-day engagement metrics for messages sent on the January 7 through March 10, 2020, mailing dates. "Click rate" is calculated using the ESD filer login page link, which was the most-clicked link and provides a lower bound on the number of unique users who followed a link from the treatment email.

---

[20] For example, the user engagement system could report that 500 unique users clicked on Link A and 400 unique users clicked on Link B, but it could not report how many clicked on multiple links because it did not record the overlap between the two groups. As a result, although we cannot calculate, with complete certainty, the number of treatment group members induced to click on at least one link, the largest of the numbers acts as a logical lower bound: If we assume that all the 400 who clicked on Link B also clicked on Link A, then 500 email recipients clicked at least one link (and 400 clicked two links). Likewise, the sum of the numbers acts as a logical upper bound on the number of treatment group members who clicked any link: If we assume there is no overlap between the groups who clicked on Link A and Link B, then 900 email recipients clicked at least one link (and, under these assumptions, each clicked only one). We use the lower bound to avoid overstating the effectiveness of the treatment email in generating click-throughs.

Exhibit B.5. Share of recipients who opened each of the four most popular links

These engagement patterns compare with email opening and click-through rates observed in a previous behavioral experiment at the Department of Veterans Affairs, in which emails to veterans highlighting their benefit eligibility saw open rates of just over 40 percent and click-through rates to the benefits application web page around 4 percent (SBST 2015).

2.   Impact findings

As described in Section IV.E, our main analysis found no treatment effects for any of the primary or secondary outcomes of interest. Despite the lack of impact on these top-line results, we conducted a number of secondary analyses to help check our results, as described in Section D of Appendix B above.

First, we analyzed the treatment impact on underlying exploratory variables, which might have highlighted the mechanism by which claimants responded to the treatment email. As shown in Exhibit B.6, we found no statistically significant changes in either the likelihood that claimants applied for standby status or switched from Internet to phone filing after receiving the treatment email or in the total amount of benefits they received.

Exhibit B.6. OLS impacts on underlying outcomes related to work search

| | Regression-adjusted group mean or percentage | | Difference | | |
	Treatment	Control	Percentage points/dollars	Standard error	*p*-value
Exploratory outcomes					
Applied for standby status post-sample entry (%)	6.3	6.1	0.2	0.300	0.48
Switched to phone filing (%)	13.0	13.1	-0.1	0.462	0.88
Total benefits paid (dollars)	3,607	3,593	14.6	30.3	0.63

Source: ESD claims adjudication data.

Note: Each regression controlled for county indicators, pre-separation job indicators, waiting week associated with claim, and Spanish language preference. We limited the sample for the standby application outcome to those who did not apply for standby status before filing their waiting week claim (25,653 claimants). We limited the sample for the outcome related to switching to phone filing to those who filed their waiting week claim online (20,158 claimants). We limited the sample for the total benefits outcome to those who received any benefits (21,045 claimants). The differences between groups were not statistically significant.

Second, we reanalyzed our main outcomes using probit models—an alternate method of modeling binary outcome variables. As shown in Exhibit B.7, as with our main results, we observed no statistically significant differences in these outcomes between the treatment and control groups. A similar analysis using logit models (another analytical approach for binary outcome data—not shown) found the same patterns.

Exhibit B.7. Probit analysis of treatment impacts on work search outcomes

	Regression-adjusted group means		Probit impact units	Standard error	*p*-value
	Treatment	**Control**			
Primary outcomes					
Percentage of claimants committing a work search error	6.2	6.2	0.000	0.0240	0.998
Secondary outcomes					
Percentage of claimants receiving a payment later disqualified	5.1	4.8	0.030	0.0258	0.241
Percentage of claimants with a cleared work search issue on at least one claim	2.7	2.5	0.030	0.0319	0.340
Percentage of claimants in the JSR sample committing a work search error	14.6	13.1	0.066	0.0887	0.459

Source: ESD claims adjudication data.

Note: We derived coefficients and standard errors from a probit analysis of binary claimant-level outcomes on treatment indicators, with other regression control variables excluded. We limited the sample for JSR errors to those selected for a job search review (1,245 claimants, versus 26,967 for the other outcomes). The differences between groups were not statistically significant.

Finally, in our third robustness check—grouping claimants by filing method—we allowed both control group outcome probabilities and treatment effects to vary, depending on how individuals filed their weekly claims. As shown in Exhibit B.8, about two-thirds of claimants exclusively filed online, and about 18 percent exclusively filed by phone, with the rest using both methods. (For a small number of claimants, we lacked information on their filing method; we excluded them from this exhibit and analysis.)

Exhibit B.8. Washington UI claimants, by filing method

Filing method	Number of claimants	Share of analytic sample
Internet only	17,531	67%
Telephone only	4,769	18%
Both methods	3,878	15%

Source: ESD claims data.

We found significant differences in the likelihood of the four outcomes related to work search errors for members of different filing method groups. (See the "Phone only" and "Both methods" control group columns of Exhibit B.9. These point estimates are estimated using only control group members and thus cannot result from the intervention. Instead, they indicate underlying differences between filing method subgroups, which could result if, for instance, some unobserved claimant characteristics helped determine both the likelihood of making a work search error and the choice of filing method.) However, as can be seen in the filing method-specific treatment effect columns of Exhibit B.9, we generally did not observe any statistically significant treatment effects for filers of any type, indicating that there was no differential impact of the intervention across these three groups for the outcomes we considered. Although the treatment had a marginally significant impact among phone-only filers selected for a JSR, we saw no other indicators of treatment effects either within that group or across the other groups.

Exhibit B.9. OLS analysis of treatment impacts, separately by filing method

	Control group prevalence by filing method			Treatment effects by filing method		
	Internet-only (base group outcome rate)	Phone-only (ppt difference)	Both methods (ppt difference)	Internet-only treatment effect	Phone-only treatment effect	Both methods treatment effect
Primary outcomes						
Committing a work search error (%)	7.4	-4.9***	-1.2*	-0.1	0.3	-0.4
Secondary outcomes						
Receiving a payment later disqualified (%)	4.4	0.9*	2.5***	0.1	0.5	1.1
Cleared work search issue on at least one claim (%)	1.8	2.5***	2.2***	0.1	-0.3	0.6
Committing a work search error (JSR sample, %)	4.8	20.0***	16.1***	1.2	7.9*	0.2

Source: ESD claims adjudication data.

Note: ***/**/* Statistically significant at the 1/5/10 percent level. Each regression controlled for county indicators, pre-separation job indicators, waiting week associated with claim, and Spanish language preference. The reference group was the control group of Internet-only filers. Regression intercepts and treatment effects varied by the claimant's filing group—phone only, online only, or both methods. We excluded from the analysis 789 claimants for whom filing method data were missing. We limited the sample for JSR errors to those selected for a job search review with nonmissing filing method data (1,213 claimants).