

# Nudging Timely Wage Reporting: Field Experimental Evidence from the U.S. Supplemental Security Income Program

C. Yiwei Zhang,<sup>a,\*</sup> Jeffrey Hemmeter,<sup>b</sup> Judd B. Kessler,<sup>c,d</sup> Robert D. Metcalfe,<sup>d,e</sup> Robert Weathers<sup>f</sup>

<sup>a</sup>Department of Consumer Science, University of Wisconsin-Madison, Madison, Wisconsin 53706; <sup>b</sup>Office of Research, Demonstration, and Employment Support, Social Security Administration, Baltimore, Maryland 21235; <sup>c</sup>Business Economics and Public Policy Department, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104; <sup>d</sup>National Bureau of Economic Research, Cambridge, Massachusetts 02138; <sup>e</sup>Department of Economics, University of Southern California, Los Angeles, California 90089; <sup>f</sup>Office of Retirement and Disability Policy, Social Security Administration, Baltimore, Maryland 21235

\*Corresponding author

Contact: cyzhang@wisc.edu,  <https://orcid.org/0000-0001-5265-6042> (CYZ); jeffrey.hemmeter@ssa.gov (JH); judd.kessler@wharton.upenn.edu,  <https://orcid.org/0000-0001-6411-7177> (JBK); robert.metcalfe@usc.edu (RDM); robert.weathers@ssa.gov (RW)

Received: August 31, 2021

Revised: July 20, 2022

Accepted: October 15, 2022

Published Online in Articles in Advance:  
December 30, 2022

<https://doi.org/10.1287/mnsc.2022.4645>

Copyright: © 2022 INFORMS

**Abstract.** We study a large-scale ( $n = 50,000$ ) natural field experiment implemented by the U.S. Social Security Administration aimed at increasing the timely and accurate self-reporting of wages by Supplemental Security Income (SSI) recipients. A letter reminding SSI recipients of their wage reporting responsibilities significantly increased both the likelihood of reporting any earnings and the total earnings reported. However, the specific letter content—providing social information or highlighting the salience of penalties—had no systematic effect. We develop a conservative estimate that the letters generated roughly \$5.91 in savings per dollar spent, highlighting the value of such a nudge in this important context.

**History:** Accepted by Yan Chen, behavioral economics and decision analysis.

**Supplemental Material:** The data files and online appendix are available at <https://doi.org/10.1287/mnsc.2022.4645>.

**Keywords:** nudges • field experiment • public policy • wage reporting • disability and work

## 1. Introduction

Federal means-tested programs provide resources to individuals with low income and few assets. Such programs accounted for more than \$773 billion in federal spending in 2019.<sup>1</sup> When cash assistance or in-kind transfers are provided on a recurring basis, these programs rely on the accurate and timely self-reporting of any changes in income or other resources to ensure that benefit payments are appropriately determined. Inaccurate and untimely reporting by beneficiaries presents a significant and ongoing policy challenge. For the Supplemental Security Income (SSI) program, the largest means-tested cash assistance program in the United States, improper benefit overpayments totaled more than \$4.6 billion in fiscal year 2019 alone (Social Security Administration 2020a).<sup>2</sup>

In recent years, the U.S. federal government has noted the need to reduce improper payments, particularly among programs they consider high priority, such as the SSI program, given the high costs of improper payments.<sup>3</sup> For the SSI program, these costs include not just the societal costs but also the burden placed on agency employees to identify accurate information for nonreporting recipients and to recover overpayments. Inaccurate reporting can pose serious

consequences for recipients, including potential reductions in future benefit payments and financial penalties if recipients are found to have knowingly withheld information. Recipients who fail to repay an overpayment may be referred to a collections agency or have their tax refund, wages, or other benefits garnished.

Economic models have typically focused on traditional policy tools, such as financial incentives, to motivate compliance (Slemrod 2019). However, misreporting persists despite the presence of financial penalties for noncompliance, with wage misreporting a leading cause of SSI overpayments.

Given the inability of traditional policy tools to eliminate noncompliance in this setting, one might consider motivating compliance with “nudges,” subtle interventions designed to change behavior without restricting choice or changing financial incentives (Thaler and Sunstein 2008).<sup>4</sup> However, recent work has suggested that nudges may fail to move behavior as much as researchers hope or expect when implemented “at scale” (DellaVigna and Linos 2022). It is therefore an open question whether nudges—even ones that have empirical support in the prior literature—will prove robust to the important settings in which policymakers aim to use them.

In this paper, we report findings from a large-scale, randomized field experiment designed and implemented by the Social Security Administration (SSA) to test whether behavioral nudges encourage more accurate and timely reporting of changes in wages for the SSI program.<sup>5</sup> Specifically, the SSA mailed letters to SSI recipients that highlighted the need to report wage changes, whereas a randomly selected control group received no letter. As shown in Figure 1, the letters provided simple information about reporting (included on all letters). Randomly selected recipients received an additional sentence that provided either: social information on reporting behavior, heightened salience on the penalties for noncompliance, or both. The SSI program provides a particularly useful setting to study how such behavioral nudges might improve reporting compliance among beneficiaries of means-tested programs. Given the complex rules surrounding the SSI program, both eligibility and benefit payment amounts are highly sensitive to fluctuations in income or resources, so even a nudge that induces a small shift in behavior has the potential to be cost-effective. It also allows us to test the efficacy of nudges at scale.

We find that nudging SSI recipients with a reminder letter significantly increased both the likelihood of reporting any countable earned income and the total amount reported in the three months immediately following the mailing of the letter relative to the control group that did not receive a letter. Over these three months, recipients were 0.34 percentage points more likely to report any countable earnings, an increase of 35.1% on a base of 0.97 percentage points. Recipients who receive a letter report 55.1% more earnings over this period. We also provide evidence that this effect reflects a change in wage reporting behavior to the SSA, not a change in actual labor supply or earned income.

We observe only mild postintervention persistence. Convergence of the control group to the treatment groups is expected given that the SSA eventually receives and verifies earnings reports from various sources to identify unreported wages, albeit at a significant delay relative to the recipients who receive a letter. That the treatment effect dissipates over time highlights that the earnings reported in the study are earnings that the SSA eventually observes through its other channels. Nevertheless, timely wage reporting induced by the nudge significantly reduces the overpayments that accrue before earnings reports are obtained from other sources and reduces the administrative burden of determining monthly wage amounts from the quarterly or annual earnings data that SSA generally obtains to identify overpayments.

We find no effect of the behaviorally motivated messaging in the letter on the likelihood of reporting any earnings, but we find some evidence that the amount

of earnings reported is sensitive to the wording in the letter.<sup>6</sup>

These findings suggest that a letter reminding SSI recipients of their wage reporting responsibilities can meaningfully influence behavior, either by accelerating the timing of wage reporting relative to when wages would have otherwise been reported or by inducing individuals who would never have reported wages to do so. Our intervention both lowered the incidence of overpayments and reduced the significant costs associated with reconciling reported and actual earnings information.

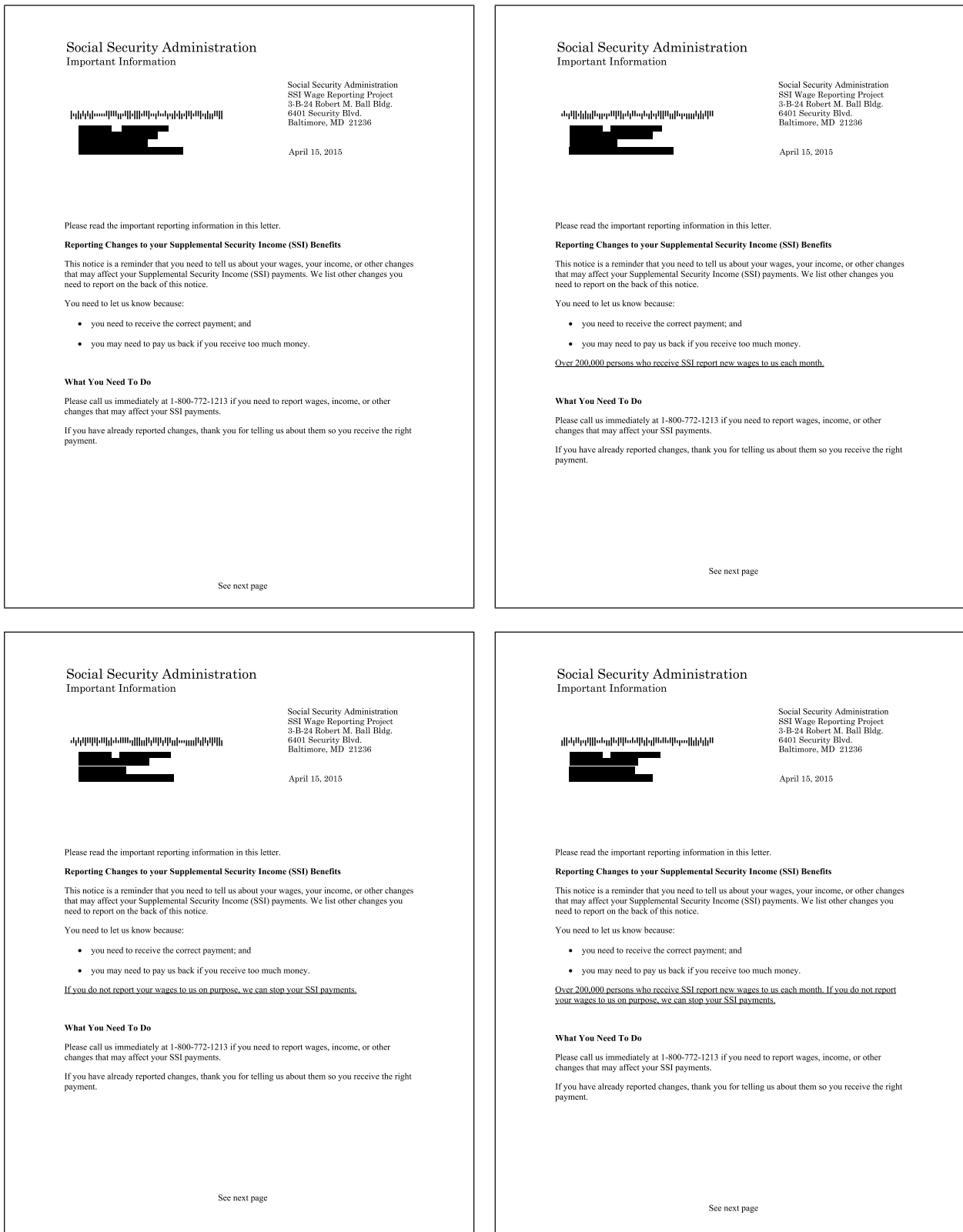
We compare the direct financial costs associated with sending the reminder letters with the costs that the SSA would have otherwise incurred for SSI recipients who failed to report their earnings in an accurate and timely manner. In doing so, we make the conservative assumption that the intervention had no positive effect on reported income beyond the first three months and that total reporting did not increase and was merely accelerated. Our conservative estimate is that the letters generated roughly \$5.91 in savings on average per dollar spent for the SSA. Although not a full social welfare evaluation, this estimate serves as a useful benchmark of the cost effectiveness of the reminder letters and suggests that not only did the reminder letters meaningfully motivate behavior, but they did so at a significant cost savings for the SSA.

Our paper highlights the importance of experimentally testing the efficacy of nudge-style interventions at scale. Although one might have worried that the reminder letters in our setting would have not been effective at encouraging reporting, we find them to be effective and highly cost-effective.

Our paper also contributes to the broader literature on the design and use of information and nudges as new policy tools available to policymakers (Bhargava and Manoli 2015, Chetty 2015) and, in particular, to the important behavioral literature demonstrating the value of reminders and informational nudges across policy domains (Frey and Meier 2004; Schultz et al. 2007; Shang and Croson 2009, Allcott 2011; Ayres et al. 2013; Allcott and Rogers 2014; Busso et al. 2015; Karlan et al. 2016a, b; Damgaard and Gravert 2018; Dai et al. 2021). Our study adds to this literature by testing the efficacy of such nudges in motivating behavior in the domain of compliance.

Most closely related to our paper is the growing literature on the use of information-based nudges to motivate tax compliance (Blumenthal et al. 2001, Fellner et al. 2013, Luttmer and Singhal 2014, Pomeranz 2015, Dwenger et al. 2016, Doerrenberg and Schmitz 2017, Hallsworth et al. 2017, Perez-Truglia and Troiano 2018, Alm 2019, Chirico et al. 2019, Cranor et al. 2020, Gillitzer and Sinning 2020, Holz et al. 2020, De Neve et al. 2021).

Figure 1. Sample Form Letter Mailings



Notes. Basic intervention mailing with no behavioral framing (top left), social information intervention mailing (top right), salience of penalties intervention mailing (bottom left), and mailing with both social information and salience of penalties language (bottom right) used in the experiment. Additional information on reporting responsibilities appeared on the back of the mailing and was identical across the four mailings.

Downloaded from informs.org by [144.92.182.142] on 14 March 2023, at 08:50 . For personal use only, all rights reserved.

In a similar study to ours, Bott et al. (2020) use a randomized field experiment to investigate the efficacy of an information letter in encouraging accurate (foreign) income reporting to the Norwegian Tax Authority. The authors find evidence that a letter making salient the moral argument for compliance and a letter aimed at increasing the perceived probability of detection both increased income reporting and differentially so, with the former having a stronger impact on the intensive margin and the latter having a stronger impact on the extensive margin. In contrast, although we find an overall effect, we find no differences by the specific behavioral framing of our reminder letters. One key feature of our setting that might explain why our findings differ from that of Bott et al. (2020) is that SSI benefits typically represent the primary, if not only, source of income for recipients. Given this, a reminder letter from the SSA, which holds authority over benefit issuance, may be sufficiently concerning that the overall effect of simply being reminded to report wages overwhelms any differential response to the message framing. More generally, it is likely that the nature of the relationship between the sender and recipient of a letter matters for how the content in the letters affect behavior. For example, recent studies in other domains have shown that the relative reputation of individuals invited to participate in a task can affect their likelihood of accepting the request (DellaVigna and Pope 2018, Chen et al. 2021). We view the interaction between the efficacy of a nudge and the relative status of the nudge sender and nudge recipient as an important area for future research.

## 2. Institutional Background

### 2.1. Background on the SSI Program

The SSI program was established in 1972 and is the largest means-tested cash assistance program in the United States, providing income support to needy blind, disabled, or aged individuals. Administered by the SSA, the program distributed nearly \$56 billion to more than 8 million eligible recipients in 2019 (Social Security Administration 2020b).

The SSI program disburses federal benefit payments to eligible recipients on a monthly basis. Both eligibility and monthly benefit amounts depend on a recipient's "countable" financial resources and "countable" income—resources and income minus any applicable exclusions—and a set of age, disability status, and residency requirements.<sup>7</sup> To be considered eligible, individuals face an asset limit of \$2,000, whereas couples face an asset limit of \$3,000.<sup>8</sup> In 2015, when our intervention took place, those who were deemed eligible could then receive up to a maximum federal benefit of \$733 per month for an individual and \$1,100 per month

for a couple. The first \$20 of monthly unearned income from nearly any source and the first \$65 of monthly earned income plus one-half remaining earnings are excluded from an individual's countable income.<sup>9</sup>

SSI recipients are required to report to the SSA any changes in circumstances that could affect their program eligibility or monthly benefit amount. Failure to report these changes as they arise can result in improper payments to recipients who are either no longer eligible or receive benefits in excess of their eligibility. Although the overall rate of payment errors is low, given the size of the SSI program and the significant dollar amounts associated with payments, even a small percentage of payment errors can add up to significant program costs. In fiscal year 2019, payment errors resulted in an estimated \$4.6 billion in overpayments, or roughly 8% of total outlays for the SSI program (Social Security Administration 2020a).

### 2.2. Wage Reporting

A major cause of improper SSI payments is recipients' failure to report new or increased wages in an accurate and timely manner. A recipient can make an initial self-report of their earnings information by contacting an SSA teleservice center or by reporting directly to one of 1,235 SSA field offices in-person, by telephone, by mail, or by fax. Most recipients can then enroll in one of two services—an automated telephone wage reporting service or a free mobile wage reporting smartphone application—to regularly report monthly wages that will affect their payment amount.<sup>10</sup>

In the absence of self-reporting, a recipient's wages must be determined through other means. The SSA uses either quarterly earnings data from the Office of Child Support Enforcement (OCSE) or annual earnings data from the Internal Revenue Service (IRS) to generate leads on possible wage activity. Although the SSA can identify unreported wages using these two sources, it can generally only do so at a considerable delay from when an overpayment was made. This delay is a result of both the time it takes for earnings data from the OCSE or IRS to be transmitted to the SSA and the time it takes for an SSA employee to fully develop and process the case before the wages can be posted to the SSA's primary administrative data file used for determining payment amounts. This latter task of identifying accurate wage information is especially time-consuming and places a considerable burden on SSA employees because they must contact the recipient and obtain primary evidence on their monthly earnings. The complicated nature of this process can be seen in the detailed policy instructions for monthly wage verification that SSA provides in its Program Operations Manual System.<sup>11</sup> This process may entail multiple attempts to contact the recipient, waiting for the recipient to submit primary evidence, identifying



whether the recipient is eligible for work incentives, and due process notification. In recognition of the significant burden, SSA emphasizes to employees who perform wage verification that the best way to avoid the extra work associated with processing wages is to process wages as soon as possible.

### 3. Experimental Design and Data

#### 3.1. Experimental Design

We study a large natural field experiment designed and implemented by the SSA to evaluate the effectiveness of several behavioral nudges on wage reporting behavior by SSI recipients.<sup>12</sup> In April 2015, the SSA mailed a sample of SSI recipients wage reporting reminder letters. Randomly selected subjects either received a basic reminder letter or a letter with an additional line of text noting other recipients' wage reporting behavior (social information), the economic penalties associated with failing to report (saliency of information), or both.<sup>13</sup> A control group received no letter. Specifically, 50,000 SSI recipients were randomly assigned to one of five groups:

1. *Basic Intervention*—a group that was sent a form letter reminding them of the need to report any information about earnings that might affect SSI payment amounts without our behavioral framing;

2. *Social Information Intervention*—a group that was sent a form letter identical to that of the “Basic Intervention” group but included additional information on the overall reporting behavior of SSI recipients (“Over 200,000 persons who receive SSI report new wages to us each month.”)<sup>14</sup>;

3. *Saliency of Penalties Intervention*—a group that was sent a form letter identical to that of the “Basic Intervention” group but included additional information on the possible financial penalties that could be incurred if wages are unreported (“If you do not report your wages to us on purpose, we can stop your SSI payments.”);

4. *Both Interventions*—a group that was sent a form letter identical to that of the “Basic Intervention” group but both contained peer information and made salient the potential financial penalties from failure to report (“Over 200,000 persons who receive SSI report new wages to us each month. If you do not report your wages to us on purpose, we can stop your SSI payments.”); and

5. *Control*—a group that was not sent a letter and therefore served as a control condition.<sup>15</sup>

As Figure 1 shows, the four experimental mailings were identical to one another and varied only in their potential inclusion of the additional lines of text.<sup>16</sup>

This intervention draws on and expands a rich behavioral literature demonstrating the value of reminders in motivating individual behavior in a wide range of

contexts.<sup>17</sup> The two types of behavioral messaging we consider—social information and saliency of information—are of particular interest in this policy setting because of their demonstrated efficacy in nudging behavior, especially among low-income individuals, the population most likely to be targeted by federal means-tested assistance programs like the SSI.<sup>18</sup>

#### 3.2. Data

Our primary source of data are the Supplemental Security Record (SSR) master file, an administrative data file containing the universe of SSI recipients. The SSR file provides demographic information (e.g., age, race, sex, state of residence, primary spoken language, and institutionalization status), program participation information (e.g., time as recipient and whether the recipient has a representative payee), and a detailed history of monthly benefit payments, beginning from the inception of the SSI program. Importantly, the SSR includes detailed information on the earned income for a recipient, allowing us to examine two key outcomes: whether recipients reported any countable earnings for themselves and the dollar amount of countable earnings reported.

Two important features of the data are worth noting. First, the extracts of the SSR we access are updated monthly, so an extract from a particular month provides a snapshot of the SSR data as it exists near the end of that month. Changes in reported earnings information are therefore observed on a monthly basis. Second, SSI recipients are generally able to report changes not only for the current month but both prospectively and retroactively as well. For instance, a recipient could revise their earnings information for the two months prior, even if they have already received SSI payments for those past two months. Thus, any observed response to the intervention could be driven by changes in reported past or future earnings. We further expand on the implications of these features of the data in our discussion of the results.

#### 3.3. Sample Selection

For the experiment, the SSA first identified a target population of approximately 240,000 SSI recipients who met the following conditions in 2015: between 18 and 50 years of age, with English as their primary language, living in the 50 states or Washington, DC, not currently institutionalized, no representative payee acting on their behalf, an SSI recipient for less than 6 years,<sup>19</sup> currently receiving payments, and with no countable earned income posted on the SSR in March 2015.<sup>20</sup> Because the proposed interventions were aimed at increasing compliance in wage reporting, these restrictions allowed for a focus on a target population currently without earnings but who were relatively likely to experience changes in their earnings.

From this target population, the SSA selected a final sample of 50,000 recipients based on an internal scoring model that prioritizes recipients for annual “redetermination” of benefit eligibility based on their likelihood of having had a change in circumstances that would affect their payment amount.<sup>21</sup> To avoid any potential confounding effects, the SSA excluded from consideration any individual whose predicted score led to being selected for redetermination (about 2.8 million SSI recipients). The final sample consists of the 50,000 highest scoring individuals (i.e., those most likely to have had a change of circumstances affecting payment amount) from the target population who were not scheduled for redetermination.

Table 1 presents summary statistics for the universe of working-age SSI recipients in current pay status as of March 2015 (column 1) and the experimental sample (column 2). The differences between the two groups are largely as would be expected given the selection criteria. The experimental sample is younger, has received SSI benefits for fewer years on average, and is more likely to be classified by the SSA as a case where a medical improvement is eventually expected. The experimental sample also did not have any reported earnings in March 2015, as expected from the sample

selection criteria, although few SSI recipients (only 2.96% of the full universe of recipients in current payment status at the time of selection) have any reported earnings. Full details on the sample characteristics can be found in Table A.1 in the online appendix. In general, the experimental sample is more similar to the target population than the universe of working age SSI recipients.

## 4. Results

We are interested in the effect of being sent a reminder letter on the likelihood of reporting any (countable) earned income and the total amount reported and any differential effect by the additional line of text. As noted in Section 3.2, two features of the data have important implications for our discussion of the results. First, an extract of the data from a particular month provides a snapshot of the SSR data as they exist near the end of that month. Second, SSI recipients can report changes both prospectively and retroactively, so every month represents a new opportunity for recipients to update their past or future earnings information. When presenting our results, we therefore report estimates (1) for a particular extract of the SSR data from a given point in time (e.g., from the July 2015 extract)

**Table 1.** Summary Statistics

|  | (1)<br>Universe of working-age<br>SSI recipients | (2)<br>Experimental<br>sample |
|--|--|-------------------------------|
| Panel A: Overall                               |  |                               |
| Reported any countable earnings in March 2015  | 2.96   | 0.00                          |
| Countable earnings reported in March 2015 (\$) | 5.53   | 0.00                          |
| Time on SSI program (yr)                       | 11.12  | 2.97                          |
| Panel B: Demographic information               |  |                               |
| Age  | 47.18  | 35.91                         |
| Race   |  |                               |
| Asian  | 1.83   | 0.64                          |
| Black  | 23.12  | 17.62                         |
| Hispanic                                       | 8.21   | 4.38                          |
| Native American or Alaskan Native              | 0.87   | 0.62                          |
| Other/unknown                                  | 26.41  | 43.89                         |
| White  | 39.56  | 32.86                         |
| Female   | 54.29  | 52.13                         |
| Has English as a primary language              | 87.67  | 100.00                        |
| Panel C: Disability information                |  |                               |
| Has a permanent disability                     | 29.39  | 10.62                         |
| Medical diary type                             |  |                               |
| Medical improvement expected (MIE)             | 2.73   | 6.62                          |
| Medical improvement possible (MIP)             | 67.88  | 82.77                         |
| Medical improvement not expected (MINE)        | 29.39  | 10.62                         |
| Observations                                   | 4,163,718  | 50,000                        |

*Notes.* This table provides summary statistics for the universe of working-age (age 18–64) SSI recipients receiving cash benefits for a disability or blindness who received a payment in March 2015 (column 1) and the subset of recipients in our experimental sample (column 2) based on data from the SSA Supplemental Security Record. Panel A reports overview statistics, Panel B reports basic demographic information, and Panel C reports details on disability status and medical diary type across each sample. All table entries represent group means. The count of individuals in each group is listed in the final row.

and (2) for income earned within a particular time horizon (e.g., income earned during the three months after intervention). Reporting the results in this way allows us to observe various potential changes in reported income induced by the letter.

#### 4.1. Effects on the Likelihood of Reporting Any Countable Earned Income

Table 2 reports estimates of the effect of receiving the letter on the likelihood of reporting any countable earnings. Each panel presents estimates using a snapshot of

the SSR data from a different point in time. Each column is a separate ordinary least squares (OLS) regression, with each pair of columns reporting the estimated effect (without and with controls) of the reminder letters on the likelihood of reporting having earned any income during the time horizon given in the heading of those two columns. The estimate for “Constant” gives the baseline likelihood of reporting for the control group of SSI recipients who did not receive any letter (for the regressions without controls), and the estimate for the indicator “Received Letter” represents the relative effect

**Table 2.** Reported Any Countable Earnings

|   | January to April<br>2015 |                       | May to July<br>2015   |                       | May to October<br>2015 |                     | May to December<br>2015 |                    | January to December<br>2015 |                     |
|---|--------------------------|-----------------------|-----------------------|-----------------------|------------------------|---------------------|-------------------------|--------------------|-----------------------------|---------------------|
|   | (1)                      | (2)                   | (3)                   | (4)                   | (5)                    | (6)                 | (7)                     | (8)                | (9)                         | (10)                |
| Panel A: SSR monthly extract from April 2015    |                          |                       |                       |                       |                        |                     |                         |                    |                             |                     |
| Received letter                                 | 0.0016**<br>(0.0008)     | 0.0016*<br>(0.0008)   |                       |                       |                        |                     |                         |                    |                             |                     |
| Constant  | 0.0041***<br>(0.0007)    |                       |                       |                       |                        |                     |                         |                    |                             |                     |
| R <sup>2</sup>                                  | 0.000                    | 0.006                 |                       |                       |                        |                     |                         |                    |                             |                     |
| p-value of F-test                               | 0.615                    | 0.592                 |                       |                       |                        |                     |                         |                    |                             |                     |
| Panel B: SSR monthly extract from July 2015     |                          |                       |                       |                       |                        |                     |                         |                    |                             |                     |
| Received letter                                 | 0.0035***<br>(0.0012)    | 0.0036***<br>(0.0012) | 0.0033***<br>(0.0012) | 0.0034***<br>(0.0012) |                        |                     |                         |                    |                             |                     |
| Constant  | 0.0087***<br>(0.0011)    |                       | 0.0097***<br>(0.0011) |                       |                        |                     |                         |                    |                             |                     |
| R <sup>2</sup>                                  | 0.000                    | 0.009                 | 0.000                 | 0.009                 |                        |                     |                         |                    |                             |                     |
| p-value of F-test                               | 0.930                    | 0.904                 | 0.916                 | 0.905                 |                        |                     |                         |                    |                             |                     |
| Panel C: SSR monthly extract from October 2015  |                          |                       |                       |                       |                        |                     |                         |                    |                             |                     |
| Received letter                                 | 0.0017<br>(0.0014)       | 0.0018<br>(0.0013)    | 0.0027*<br>(0.0014)   | 0.0028*<br>(0.0014)   | 0.0029*<br>(0.0016)    | 0.0030*<br>(0.0016) |                         |                    |                             |                     |
| Constant  | 0.0135***<br>(0.0012)    |                       | 0.0146***<br>(0.0013) |                       | 0.0185***<br>(0.0014)  |                     |                         |                    |                             |                     |
| R <sup>2</sup>                                  | 0.000                    | 0.011                 | 0.000                 | 0.012                 | 0.000                  | 0.014               |                         |                    |                             |                     |
| p-value of F-test                               | 0.322                    | 0.257                 | 0.518                 | 0.456                 | 0.484                  | 0.420               |                         |                    |                             |                     |
| Panel D: SSR monthly extract from December 2015 |                          |                       |                       |                       |                        |                     |                         |                    |                             |                     |
| Received letter                                 | 0.0014<br>(0.0015)       | 0.0015<br>(0.0015)    | 0.0022<br>(0.0016)    | 0.0023<br>(0.0016)    | 0.0023<br>(0.0018)     | 0.0025<br>(0.0018)  | 0.0025<br>(0.0018)      | 0.0027<br>(0.0018) | 0.0032<br>(0.0020)          | 0.0034*<br>(0.0020) |
| Constant  | 0.0173***<br>(0.0013)    |                       | 0.0183***<br>(0.0014) |                       | 0.0245***<br>(0.0016)  |                     | 0.0259***<br>(0.0016)   |                    | 0.0306***<br>(0.0018)       |                     |
| R <sup>2</sup>                                  | 0.000                    | 0.012                 | 0.000                 | 0.012                 | 0.000                  | 0.017               | 0.000                   | 0.017              | 0.000                       | 0.019               |
| p-value of F-test                               | 0.230                    | 0.178                 | 0.513                 | 0.442                 | 0.297                  | 0.245               | 0.498                   | 0.435              | 0.275                       | 0.216               |
| Controls  |                          | Yes                   |                       | Yes                   |                        | Yes                 |                         | Yes                |                             | Yes                 |
| Observations                                    | 50,000                   | 50,000                | 50,000                | 50,000                | 50,000                 | 50,000              | 50,000                  | 50,000             | 50,000                      | 50,000              |

*Notes.* This table reports estimates of the aggregate effect of receiving a letter reminding SSI recipients of their wage reporting responsibilities on the likelihood of reporting any countable earnings. Each column presents a separate regression where the dependent variable is an indicator for whether the recipient reported any countable earnings. Controls include age, gender, dummies for race/ethnicity, state or territory of residence, years on the SSI program, an indicator for whether the recipient is permanently disabled, and indicators for the primary disability diagnosis. The time horizon for the dependent variable varies by column. Columns 1 and 2 report estimates where the time horizon for the dependent variable is the three months prior to and including the month when the treatment letters were mailed (i.e., whether any countable earnings were reported for January 2015–April 2015); Columns 3 and 4 report estimates where the time horizon is the three months after treatment (May 2015–July 2015); columns 5 and 6 report estimates where the time horizon is the six months after treatment (May 2015–October 2015); columns 7 and 8 report estimates where the time horizon is after treatment through the year end (May 2015–December 2015); and columns 9 and 10 report estimates where the time horizon is the full calendar year for 2015. Each panel presents estimates using a different extract of the SSR as it existed at the end of the indicated month. For example, Panel A reports estimates using an extract of the SSR data as it existed at the end of April 2015. Coefficients are reported for a “Received letter” indicator denoting whether the recipient received a letter (experimental mailing) from the SSA. The last row of each panel reports *p* values from an *F* test of whether there is no differential effect of the behavioral framing.

\*, \*\*, and \*\*\*Significance levels at the 10%, 5%, and 1% levels, respectively.

Downloaded from informas.org by [144.92.182.142] on 14 March 2023, at 08:50 . For personal use only, all rights reserved.

of having received any of the four versions of the letter on the likelihood of reporting.

Columns 1 and 2 of Panel A show that by the end of April 2015 (approximately two weeks after mailing), having received a letter is associated with a moderately significant 0.16-percentage-point increase ( $p < 0.10$ ) in the likelihood of reporting any countable earned income during the months January to April of 2015. This represents a 39.0% increase in the likelihood of reporting on a mean likelihood of 0.41 percentage points in the control group and implies that receiving the letter induced treated individuals to contact SSA to report earnings from those four months.<sup>22</sup> By the end of July 2015 (Panel B), three months after intervention, the estimated effect on reporting any countable earned income from these four months increases in both magnitude and significance. By the end of the calendar year (Panel D), the effect has decayed somewhat and is no longer significant.

Columns 3 and 4 show a similar pattern on reporting any countable earned income from the first three months after intervention (May to July 2015). By July 2015 (Panel B), having received a letter is associated with a 0.34-percentage-point increase ( $p < 0.01$ ) in the likelihood of reporting any countable earned income during those three months (May to July 2015). This represents a 35.1% increase in the likelihood of reporting on a mean likelihood of 0.97 percentage points in the control group. This effect also decreases in both magnitude and significance by the calendar year end. Looking at the baseline likelihood of reporting for the control group, the decay in the estimated treatment effect is driven by the control group eventually catching up to the treatment groups in the likelihood that the SSA has a record of countable earned income—albeit at a delay relative to the recipients who received a letter.<sup>23</sup>

Looking over the full calendar year, columns 9 and 10 of Panel D report the estimated effect of receiving a letter on the likelihood of reporting any countable earned income during 2015 by the end of the year. SSI recipients who received a letter were on average 0.33 percentage points ( $p < 0.10$ ) more likely to have reported any earnings for the year. This represents an 11.1% increase in the likelihood of reporting on a mean likelihood of 3.06% by recipients who did not receive a letter.<sup>24</sup> Taken together, these findings suggest that the reminder letters did meaningfully nudge SSI recipients to report changes in their countable earned income. Much of the treatment effect came in the form of reporting countable earnings earlier than they would have otherwise, although some of the effect appears to persist through the year end.<sup>25</sup>

The bottom row of each panel in Table 2 reports results from  $F$  tests of whether there was a statistically significant effect of which behavioral messaging recipients received. Although there is clearly an effect of

receiving the reminder letter, the specific behavioral messaging in the letter does not appear to have mattered. It may be that any notification from SSA is taken very seriously by this population, given their reliance on these benefits, or that the basic information included in all versions of the letter addressed behavioral biases; however, it is not possible to disentangle the reason for no observed differential impacts with the current data. Table A.6 in the online appendix presents the full results by treatment.

#### 4.2. Effects on the Total Dollar Amount of Reported Countable Earnings

Table 3 reports estimates of the effect of receiving a letter on the dollar amount of countable earnings reported. Each panel again uses a snapshot of the SSR data from a different point in time. Each column is a separate OLS regression, with each pair of columns reporting the estimated effect (without and with controls) during the time horizon given in the heading of the two columns. Reports of no earnings are included in the regressions as zeros. The estimate for “Constant” gives the baseline amount of reported earnings for the control group (for the regressions without controls), and the estimate for the indicator “Received Letter” represents the relative effect of having received any of the four versions of the letter on the amount of reported earnings.

We find similar patterns of behavior in Table 3 as were reported in Table 2. There is an initial significant effect of the reminder letters. Two weeks after mailing, SSI recipients who received a letter have reported countable earnings during the months of January to April 2015 that are \$1.96 higher on average ( $p < 0.10$ ) than the average reported countable earnings of the control group (\$2.67 over the same time horizon), representing a 73.4% increase in reported earnings. By July 2015, SSI recipients who received a letter have reported countable income earned during the three months after intervention (May to July of 2015) that is \$4.89 higher on average ( $p < 0.01$ ) than the average reported countable earnings of the control group (\$8.88 over the same time horizon), representing a 55.1% increase in reported countable earnings. As with the extensive margin results, this immediate effect decreases in both magnitude and significance by year end. Looking over the full calendar year (columns 9 and 10 of Panel D), we see no significant difference in the amount of countable earnings reported by whether SSI recipients received a reminder letter.<sup>26</sup>

We do, however, observe statistically significant differences in the total amount of reported countable earnings for the full calendar year as a function of the behavioral messaging included in letter. As with Table 2, the bottom row of each panel in Table 3 report results from  $F$  tests of whether we can reject that the



**Table 3.** Amount of Countable Earnings Reported

|   | January to April<br>2015 |           | May to July<br>2015 |           | May to October<br>2015 |          | May to December<br>2015 |          | January to December<br>2015 |          |
|---|--------------------------|-----------|---------------------|-----------|------------------------|----------|-------------------------|----------|-----------------------------|----------|
|   | (1)                      | (2)       | (3)                 | (4)       | (5)                    | (6)      | (7)                     | (8)      | (9)                         | (10)     |
| Panel A: SSR monthly extract from April 2015    |                          |           |                     |           |                        |          |                         |          |                             |          |
| Received letter                                 | 1.9736*                  | 1.9550*   |                     |           |                        |          |                         |          |                             |          |
|   | (1.0792)                 | (1.0797)  |                     |           |                        |          |                         |          |                             |          |
| Constant  | 2.6654***                |           |                     |           |                        |          |                         |          |                             |          |
|   | (0.9652)                 |           |                     |           |                        |          |                         |          |                             |          |
| R <sup>2</sup>                                  | 0.000                    | 0.003     |                     |           |                        |          |                         |          |                             |          |
| p-value of F-test                               | 0.278                    | 0.285     |                     |           |                        |          |                         |          |                             |          |
| Panel B: SSR monthly extract from July 2015     |                          |           |                     |           |                        |          |                         |          |                             |          |
| Received letter                                 | 5.7075***                | 5.8674*** | 4.6924***           | 4.8852*** |                        |          |                         |          |                             |          |
|   | (1.6645)                 | (1.6644)  | (1.7515)            | (1.7500)  |                        |          |                         |          |                             |          |
| Constant  | 5.4218***                |           | 8.8803***           |           |                        |          |                         |          |                             |          |
|   | (1.4887)                 |           | (1.5666)            |           |                        |          |                         |          |                             |          |
| R <sup>2</sup>                                  | 0.000                    | 0.004     | 0.000               | 0.006     |                        |          |                         |          |                             |          |
| p-value of F-test                               | 0.826                    | 0.826     | 0.219               | 0.237     |                        |          |                         |          |                             |          |
| Panel C: SSR monthly extract from October 2015  |                          |           |                     |           |                        |          |                         |          |                             |          |
| Received letter                                 | 4.2045**                 | 4.3783**  | 2.5463              | 2.7410    | 3.3219                 | 3.6771   |                         |          |                             |          |
|   | (2.0730)                 | (2.0722)  | (2.2868)            | (2.2844)  | (4.4857)               | (4.4789) |                         |          |                             |          |
| Constant  | 10.8549***               |           | 15.2602***          |           | 34.3017***             |          |                         |          |                             |          |
|   | (1.8541)                 |           | (2.0454)            |           | (4.0121)               |          |                         |          |                             |          |
| R <sup>2</sup>                                  | 0.000                    | 0.004     | 0.000               | 0.006     | 0.000                  | 0.007    |                         |          |                             |          |
| p-value of F-test                               | 0.631                    | 0.610     | 0.048               | 0.053     | 0.030                  | 0.032    |                         |          |                             |          |
| Panel D: SSR monthly extract from December 2015 |                          |           |                     |           |                        |          |                         |          |                             |          |
| Received letter                                 | 4.0612*                  | 4.2021*   | 2.7821              | 2.9653    | 3.2694                 | 3.6181   | 2.7396                  | 3.2123   | 6.8007                      | 7.4144   |
|   | (2.2235)                 | (2.2221)  | (2.3815)            | (2.3785)  | (4.6433)               | (4.6349) | (6.1286)                | (6.1160) | (7.6379)                    | (7.6220) |
| Constant  | 13.8605***               |           | 17.3455***          |           | 39.3275***             |          | 55.0110***              |          | 68.8715***                  |          |
|   | (1.9888)                 |           | (2.1300)            |           | (4.1531)               |          | (5.4816)                |          | (6.8315)                    |          |
| R <sup>2</sup>                                  | 0.000                    | 0.005     | 0.000               | 0.006     | 0.000                  | 0.007    | 0.000                   | 0.008    | 0.000                       | 0.008    |
| p-value of F-test                               | 0.509                    | 0.491     | 0.089               | 0.096     | 0.041                  | 0.042    | 0.054                   | 0.055    | 0.086                       | 0.084    |
| Controls  |                          | Yes       |                     | Yes       |                        | Yes      |                         | Yes      |                             | Yes      |
| Observations                                    | 50,000                   | 50,000    | 50,000              | 50,000    | 50,000                 | 50,000   | 50,000                  | 50,000   | 50,000                      | 50,000   |

*Notes.* This table reports estimates of the aggregate effect of receiving a letter reminding SSI recipients of their wage reporting responsibilities on the dollar amount of countable earnings reported. Each column presents a separate regression where the dependent variable is the amount of countable earnings reported by the recipient. Controls include age, gender, dummies for race/ethnicity, state or territory of residence, years on the SSI program, an indicator for whether the recipient is permanently disabled, and indicators for the primary disability diagnosis. The time horizon for the dependent variable varies by column. Columns 1 and 2 report estimates where the time horizon for the dependent variable is the three months prior to and including the month when the treatment letters were mailed (i.e., whether any countable earnings were reported for January 2015–April 2015); columns 3 and 4 report estimates where the time horizon is the three months after treatment (May 2015–July 2015); columns 5 and 6 report estimates where the time horizon is the six months after treatment (May 2015–October 2015); columns 7 and 8 report estimates where the time horizon is after treatment through the year end (May 2015–December 2015); and columns 9 and 10 report estimates where the time horizon is the full calendar year for 2015. Each panel presents estimates using a different extract of the SSR data as it existed at the end of the indicated month. For example, Panel A reports estimates using an extract of the SSR data as it existed at the end of April 2015. Coefficients are reported for a “Received letter” indicator denoting whether the recipient received a letter (experimental mailing) from the SSA. The last row of each panel reports *p* values from an *F* test of whether there is no differential effect of the behavioral framing.

\*, \*\*, and \*\*\*Significance levels at the 10%, 5%, and 1% levels, respectively.

content of the letters had no effect on how much countable earnings were reported. We find somewhat differential effects on the amount of reported countable earnings depending on the behavioral messaging in the letter, with *F* test *p* values that range from 0.04 to 0.5 in Panel D of Table 3. Specifically, as Table A.7 in the online appendix shows, receiving a letter with both social information and information increasing the salience of the penalties is associated with approximately \$21.27 higher reported countable earnings

over the entire year relative to the baseline earnings for the control group that did not receive any letter.<sup>27</sup> This difference between the treatments is statistically significant in our main analysis. Unlike our other results, however, it does not survive multiple hypothesis testing corrections (List et al. 2019). We see the question of whether a combined message—which provides both social encouragement and a heightened sense of punishment—is particularly effective in other settings as an exciting avenue for future research.

### 4.3. Labor Supply Effects

One key question is whether our findings reflect changes in actual earnings. We examine the effect of our reminder letter on labor supply using the SSA Master Earnings File, which provides detailed data on annual earnings as reported to the IRS on individuals' Form W-2 or 1040 Schedule SE forms (if self-employed). Table A.8 in the online appendix shows no effect of the reminder letter on the likelihood of having been employed, the amount of W2 earnings, or the number of employers in 2015, suggesting that our findings reflect changes in reporting behavior and not increases in overall earnings or employment.

### 4.4. Cost-Effectiveness Analysis

We assess the cost-effectiveness of these reminder letters by comparing the costs associated with the letters with the costs that the SSA would have otherwise incurred for SSI recipients who failed to report their earnings in an accurate and timely manner. That is, the counterfactual we consider when evaluating the cost-effectiveness of the reminder letters is the absence of any intervention (i.e., the way the SSA traditionally identifies and recovers overpayments; Benartzi et al. 2017). Because we find no differential effects by behavioral messaging on the extensive margin, we focus our assessment on the cost-effectiveness of the reminder letters in aggregate.

The cost of printing and mailing each letter includes \$0.135 in printing costs and \$0.435 in postage costs for a total cost of \$0.57 per recipient, or \$22,800 for the 40,000 individuals (i.e., the implementation cost). We evaluate the cost-effectiveness of this intervention by using a very conservative estimate of its benefits. We focus only on statistically significant increases in aggregate income reported as of July 2015 for the months prior to and including the month of the intervention and the three months after intervention. By focusing only on this period, we make the conservative assumption that the intervention had no positive differential effect on reported income beyond the first three months.

As of July 2015, our estimates imply that the intervention led SSI recipients who received a letter to report an additional \$10.76 in countable earned income. Of this increase, an estimated \$5.87 ( $p < 0.01$ ; column 2 of Panel B in Table 3) was for income earned during the three months leading up to and including the month of the intervention and \$4.89 ( $p < 0.01$ ; column 4 of Panel B in Table 3) was for income earned during the three months after. The \$5.87 increase in reported income earned during the three months prior to the intervention represents a retrospective increase in reported earnings that were not likely reported in time to avoid overpayment of SSI benefits. If we conservatively assume that all unreported earnings are

eventually discovered by the SSA, then our intervention merely accelerated the discovery of the overpayments and so may not have generated meaningful savings. However, the additional \$4.89 in income reported during the three months postintervention led to a contemporaneous reduction in payments. Even if the earnings were later discovered by the SSA, the SSA would likely have recovered only part of these funds. A recent report by the SSA estimates that about 69% of overpayments are never recovered (Social Security Administration 2019). Thus, by shifting the reporting of this income to occur as it was earned, our intervention saved the SSA roughly \$3.37 (\$4.89 times 0.69) per recipient. Given the total cost of \$0.57 per recipient to print and mail the letters, our highly conservative estimate suggests that the letters generated approximately \$5.91 in savings on average per \$1 spent by the SSA (\$3.37/\$0.57).

This estimate is quite conservative and almost certainly underestimates the actual savings associated with the reminder letters. For one, it assumes that any increase in reported earnings induced by the intervention would have eventually been discovered by the SSA; if any of these reported earnings would not have been discovered, then the letters would have saved the SSA even more in overpayments. It also assumes that the differential treatment effect for individuals who were nudged into reporting their earnings disappeared fully after July 2015—three months after the intervention—even though the effects of receiving a letter clearly persist beyond this three-month window. Furthermore, the estimates conservatively assume that, once an overpayment has been made, when earnings are discovered does not matter. However, the SSA has a higher recovery rate when earnings are discovered earlier, suggesting benefits from the \$5.87 in countable earnings reported in the months before the intervention. These calculations also do not account for the costs associated with a scheduled redetermination that might be avoided due to these reports (approximately \$200 on average; see Social Security Administration 2015). Nor does it account for the costs to SSA employees of their time to reconcile any unreported earnings. Taken together, our rough calculations suggest that in addition to having a meaningful impact on wage reporting behavior, the reminder letters were also quite cost-effective as a behaviorally informed policy tool.

A caveat to our assessment is that it is not a full social welfare evaluation of the reminder letters. We do not consider any positive or negative utility associated with being nudged to report earnings or directly imposed by the nudge itself (Allcott and Kessler 2019, Butera et al. 2022). We also do not account for any potential indirect costs that may have resulted from increases in the use of or reliance on other sources of

support, including other welfare programs, or from any increased burden on the agency. Rather, we focus our attention on a comparison of the direct pecuniary costs and benefits associated with sending the reminder letters and the direct pecuniary costs associated with the standard process for recovering overpayments.

## 5. Conclusion

There is a growing interest among policymakers and academics in leveraging insights from psychology and behavioral economics to design nudges to encourage positive behavior, particularly in settings where traditional policy tools have not been fully effective. However, there is also a concern that such insights might fail to deliver on their potential when implemented “at scale” in relevant policy settings.

We examine the results of a large-scale, randomized field experiment aimed at encouraging the more accurate and timely reporting of changes in wages for the SSI program. We find that sending a reminder letter, regardless of whether behaviorally motivated language was included, significantly increased the likelihood that subjects reported any countable earnings and the amount of earnings reported in the three months immediately following the intervention.

The results of our study add to an important empirical literature highlighting the potential of such interventions to improve outcomes in policy settings (Hallsworth et al. 2015, 2017). Our approach of using a field experiment with granular individual data to test behavioral interventions in policymaking is in concordance with the recommendations of the recent U.S. *Foundations for Evidence-Based Policymaking Act of 2018* (Hahn 2019).<sup>28</sup> We hope that more behaviorally informed interventions and policies at the federal, state, and local levels are tested and optimized using natural field experiments to explore whether other interventions can also deliver on the promise of influencing behavior at scale in a way that is highly cost effective. Our findings also suggest the potential efficacy of simple reminders in managerial settings, where encouraging compliance with expected or required behaviors is an important challenge.

## Acknowledgments

The authors thank Maya Shankar, the White House Social and Behavioral Sciences Team, and the U.S. General Services Administration Office of Evaluation Sciences for support of this project. The opinions, conclusions, and errors in this paper are the sole responsibility of the authors and do not represent the official views of the Social Security Administration, the University of Pennsylvania, the University of Southern California, or the University of Wisconsin-Madison and do not assume endorsement by the Federal Government. The experiment in this study is registered at the American Economic Association Social Science Registry (Randomized Control Trial

ID AEARCTR-0005351, available at <http://www.socialscienceregistry.org/trials/5351>).

## Endnotes

<sup>1</sup> This figure was projected as of May 2019 (Congressional Budget Office 2019).

<sup>2</sup> The Office of Management and Budget (OMB) defines an improper payment as any payment that should not have been made or that was made in an incorrect amount under statutory, contractual, administrative, or other legally applicable requirements, including payments with insufficient documentation to determine if the payment was proper. Based on OMB guidance, any program with \$750 million in improper payments is considered a high-priority program.

<sup>3</sup> The Improper Payments Information Act of 2002, as amended by the Improper Payments Elimination and Recovery Act of 2010 and the Improper Payments Elimination and Recovery Improvement Act of 2012, requires agencies to periodically review all programs and activities. For any identified as susceptible to significant improper payments, agencies must estimate and report the amount of improper payments and detail their efforts to monitor and minimize improper payments.

<sup>4</sup> See Benartzi et al. (2017) for an overview of some key empirical studies on “nudge” interventions in policy settings and the important cost considerations necessary for evaluating their effectiveness.

<sup>5</sup> Previous research has shown how the SSI affects labor supply decisions (Neumark and Powers 2000, Kaushal 2010, Deshpande 2016) and financial well-being (Deshpande et al. 2021).

<sup>6</sup> Although not robustly statistically significant, the combined message appears relatively more effective than either of the behavioral messages on their own.

<sup>7</sup> Eligibility and payment levels also depend on spousal or parental (for children) income and on living arrangements, specifically whether the individual or couple lives in their own household, in another’s household, or in a Medicaid facility.

<sup>8</sup> Countable resources are generally defined as “cash or other liquid assets or any real or personal property that individuals (or their spouses) own and could convert to cash to be used for their support and maintenance.” Exclusions to these asset limits include, but are not limited to, the value of a recipient’s home, a vehicle used for transportation, and any household goods and personal effects. See Social Security Administration (2020b) for additional details.

<sup>9</sup> Any residual amount from the \$20 income exclusion applied to unearned income is applied to earned income instead. For example, an individual with wages of \$535 per month and no other sources of income would have \$225 in countable income (wages minus \$85 in income exclusions, reduced by half). The maximum federal benefit of \$733 per month is then reduced by her countable income, giving her \$508 per month in SSI benefits.

<sup>10</sup> An online reporting tool now exists but was not yet available in 2015 during the experiment.

<sup>11</sup> The instructions (effective as of 10/22/2020) can be found at <https://secure.ssa.gov/poms.nsf/lnx/0500820130>.

<sup>12</sup> We focus our main analysis on wage reporting rather than labor supply as this was the primary focus of the SSA. However, in Section 4.3, we show that the treatment had no effect on actual earnings or labor supply.

<sup>13</sup> As shown in Figure 1, the additional line of text providing social information or making penalties salient was included in the middle of the letter and underlined for added emphasis. Nonetheless, it is possible that recipients did not pay attention to this additional information, which could lead our tests of these treatments to be



underpowered. Prior research, however, has shown that presenting information in this way can be sufficient to generate positive treatment effects (Coffman et al. 2017, Bott et al. 2020).

<sup>14</sup> Whether individuals upwardly or downwardly revise their understanding of social norms around reporting in response to the social information intervention depends on their prior beliefs regarding SSI recipient behavior. Although we expect the information provided to lead to an upward revision in beliefs, our analysis is agnostic and allows for either increases or decreases in reporting behavior in response to the intervention.

<sup>15</sup> Because our control group receives no letter, we are not able to distinguish between the effect of receiving a letter with a reminder sentence versus simply receiving a letter from SSA (e.g., a letter informing a beneficiary of an adjustment to their benefit payment amount).

<sup>16</sup> All letters were mailed out on the same day. Balance tests in Table A.2 of the online appendix show that the five randomized groups are similar across observable characteristics. Of the 96 pretreatment characteristics, only 9 were statistically significantly different across treatment and control groups ( $p < 0.1$ ), and none of these differences were economically meaningful. Using extracts of the Supplemental Security Record from before the intervention, we also show in Table A.3 of the online appendix that being assigned to receive a letter has no effect on wage reporting behavior for the months prior to the intervention. This serves as a useful check because being randomly assigned to receive a letter in April 2015 should have no effect on whether—as of March 2015—you have reported earning any income during the three months before the intervention (January to March 2015).

<sup>17</sup> See, for example, the efficacy of reminders for medical care (Busso et al. 2015), financial behaviors (Karlan et al. 2016a, b), tax compliance (Hallsworth et al. 2017, Chirico et al. 2019), and charitable giving (Damgaard and Gravert 2018).

<sup>18</sup> See, for example, the influence of social information about others for motivating people to take a teaching job (Coffman et al. 2017), vote (Gerber and Rogers 2009, Keane and Nickerson 2015), and pay taxes promptly (Hallsworth et al. 2017). See also the efficacy of increasing the saliency of information about the Earned Income Tax Credit for reported earnings (Chetty and Saez 2013, Chetty et al. 2013, Bhargava and Manoli 2015); eligibility information for enrollment in the Supplemental Nutrition Assistance Program (Finkelstein and Notowidigdo 2019); and information on social security for the decision to work (Liebman and Luttmer 2015).

<sup>19</sup> Newer SSI recipients are more likely to work than longer-term recipients (Ben-Shalom and Stapleton 2015).

<sup>20</sup> Most SSI recipients have zero earnings.

<sup>21</sup> As part of a redetermination, the SSA reviews the income, resources, and living arrangements of a recipient to determine whether they are still eligible and receiving the correct payment amount.

<sup>22</sup> The control group mean is the constant from the regression without controls (column 1 of Panel A in Table 2).

<sup>23</sup> The increase in the control group likely comes from a combination of self-reporting and wage identification through other channels. We are not able to identify the specific channel through which this increase occurs with our data.

<sup>24</sup> This low likelihood for the control group of reporting any earnings by the calendar year end (3.06%) is consistent with the typically low reporting rates for working-age SSI recipients and reflects the fact that most recipients are not working and therefore would not be expected to report.

<sup>25</sup> Table A.4 in the online appendix shows that our postintervention results in Table 2 are robust to measuring our outcomes over individual months.

<sup>26</sup> Table A.5 in the online appendix shows that, like our extensive margin results, our postintervention results in Table 3 are also robust to measuring our outcomes over individual months.

<sup>27</sup> This estimate is constructed by adding the four individual coefficients reported in column 10 of Panel D of Table A.7 in the online appendix.

<sup>28</sup> See <http://www.congress.gov/bill/115th-congress/house-bill/4174> for details on the bill.

## References

- Allcott H (2011) Social norms and energy conservation. *J. Public Econom.* 95(9–10):1082–1095.
- Allcott H, Kessler JB (2019) The welfare effects of nudges: A case study of energy use social comparisons. *Amer. Econom. J. Appl. Econom.* 11(1):236–276.
- Allcott H, Rogers T (2014) The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *Amer. Econom. Rev.* 104(10):3003–3037.
- Alm J (2019) What motivates tax compliance. *J. Econom. Survey* 33(2):351–726.
- Ayres I, Raseman S, Shih A (2013) Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *J. Law Econom. Organ.* 29(5):992–1022.
- Ben-Shalom Y, Stapleton DC (2015) Long-term work activity and use of employment supports among new supplemental security income recipients. *Soc. Security Bull.* 75(1):73–95.
- Benartzi S, Beshears J, Milkman KL, Sunstein CR, Thaler RH, Shankar M, Tucker-Ray W, et al. (2017) Should governments invest more in nudging? *Psych. Sci.* 28(8):1041–1055.
- Bhargava S, Manoli D (2015) Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment. *Amer. Econom. Rev.* 105(11):3489–3529.
- Blumenthal M, Christian C, Slemrod J (2001) Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota. *National Tax J.* 54(1):125–138.
- Bott KM, Cappelen AW, Sørensen EØ, Tungodden B (2020) You've got mail: A randomized field experiment on tax evasion. *Management Sci.* 66(7):2801–3294.
- Busso M, Cristia J, Humpage S (2015) Did you get your shots? Experimental evidence on the role of reminders. *J. Health Econom.* 44:226–237.
- Butera L, Metcalfe RD, Morrison W, Taubinsky D (2022) Measuring the welfare effects of shame and pride. *Amer. Econom. Rev.* 112(1):122–168.
- Chen Y, Farzan R, Kraut R, Zaare IY, Zhang AF (2021) Motivating experts to contribute to digital public goods: A personalized field experiment on Wikipedia. Working paper, University of Michigan, Ann Arbor, MI.
- Chetty R (2015) Behavioral economics and public policy: A pragmatic perspective. *Amer. Econom. Rev. P & P* 105(5):1–33.
- Chetty R, Saez E (2013) Teaching the tax code: Earnings responses to an experiment with EITC recipients. *Amer. Econom. J. Appl. Econom.* 5(1):1–31.
- Chetty R, Friedman J, Saez E (2013) Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. *Amer. Econom. Rev.* 103(7):2683–2721.
- Chirico M, Inman R, Loeffler C, MacDonald J, Sieg H (2019) Detering tax delinquency in Philadelphia: An experimental evaluation of nudge strategies. *National Tax J.* 72(3): 479–506.
- Coffman LC, Featherstone CR, Kessler JB (2017) Can social information affect what job you choose and keep? *Amer. Econom. J. Appl. Econom.* 9(1):96–117.
- Congressional Budget Office (2019) Federal mandatory spending for means-tested programs, 2009 to 2029. CBO Publication No. 55347, Congressional Budget Office, Washington, DC.



- Cranor T, Goldin J, Homonoff T, Moore L (2020) Communicating tax penalties to delinquent taxpayers: Evidence from a field experiment. *National Tax J.* 73(2):331–360.
- Dai H, Saccardo S, Han MA, Roh L, Raja N, Vangala S, Modi H, et al. (2021) Behavioural nudges increase COVID-19 vaccinations. *Nature* 597:404–409.
- Damgaard MT, Gravert C (2018) The hidden costs of nudging: Experimental evidence from reminders in fundraising. *J. Public Econom.* 157:15–26.
- De Neve J-E, Imbert C, Spinnewijn J, Tsankova T, Luts M (2021) How to improve tax compliance? Evidence from population-wide experiments in Belgium. *J. Political Econom.* 129(5):1425–1463.
- DellaVigna S, Linos E (2022) RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica* 90(1):81–116.
- DellaVigna S, Pope D (2018) What motivates effort? Evidence and expert forecasts. *Rev. Econom. Stud.* 85(2):1029–1069.
- Deshpande M (2016) Does welfare inhibit success? The long-term effects of removing low-income youth from the disability rolls. *Amer. Econom. Rev.* 106(11):3300–3330.
- Deshpande M, Gross T, Su Y (2021) Disability and distress: The effect of disability programs on financial outcomes. *Amer. Econom. J. Appl. Econom.* 13(2):151–178.
- Doerrenberg P, Schmitz J (2017) Tax compliance and information provision: A field experiment with small firms. *J. Behav. Econom. Policy* 1(1):47–54.
- Dwenger N, Kleven H, Rasul I, Rincke J (2016) Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany. *Amer. Econom. J. Appl. Econom.* 8(3):203–232.
- Fellner G, Sausgruber R, Traxler C (2013) Testing enforcement strategies in the field: Threat, moral appeal and social information. *J. Eur. Econom. Assoc.* 11(3):634–660.
- Finkelstein A, Notowidigdo MJ (2019) Take-up and targeting: Experimental evidence from SNAP. *Quart. J. Econom.* 134(3):1505–1556.
- Frey BS, Meier S (2004) Social comparisons and pro-social behavior: Testing “conditional cooperation” in a field experiment. *Amer. Econom. Rev.* 94(5):1717–1722.
- Gerber AS, Rogers T (2009) Descriptive social norms and motivation to vote: Everybody’s voting and so should you. *J. Politics* 71(1):178–191.
- Gillitzer C, Sinning M (2020) Nudging businesses to pay their taxes: Does timing matter? *J. Econom. Behav. Organ.* 169:284–300.
- Hahn R (2019) Building upon foundations for evidence-based policy. *Science* 364(6440):534–535.
- Hallsworth M, List JA, Metcalfe RD, Vlaev I (2015) The making of Homo honoratus: From omission to commission. NBER Working Paper No. 21210, National Bureau of Economic Research, Cambridge, MA.
- Hallsworth M, List JA, Metcalfe RD, Vlaev I (2017) The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *J. Public Econom.* 148:14–31.
- Holz JE, List JA, Zentner A, Cardoza M, Zentner J (2020) The \$100 million nudge: Increasing tax compliance of businesses and the self-employed using a natural field experiment. NBER Working Paper No. 27666, National Bureau of Economic Research, Cambridge, MA.
- Karlan D, Morten M, Zinman J (2016a) A personal touch in text messaging can improve microloan repayment. *Behav. Sci. Policy* 1(2):25–31.
- Karlan D, McConnell M, Mullainathan S, Zinman J (2016b) Getting to the top of mind: How reminders increase saving. *Management Sci.* 62(12):3393–3411.
- Kaushal N (2010) Elderly immigrants’ labor supply response to supplemental security income. *J. Policy Anal. Management* 29(1):137–162.
- Keane LD, Nickerson DW (2015) When reports depress rather than inspire: A field experiment using age cohorts as reference groups. *J. Political Marketing* 14(4):381–390.
- Liebman JB, Luttmer EFP (2015) Would people behave differently if they better understood Social Security? Evidence from a field experiment. *Amer. Econom. J. Appl. Econom.* 7(1):275–299.
- List JA, Shaikh AM, Xu Y (2019) Multiple hypothesis testing in experimental economics. *Experiment. Econom.* 22:773–793.
- Luttmer EFP, Singhal M (2014) Tax morale. *J. Econom. Perspective* 28(4):149–168.
- Neumark D, Powers E (2000) Welfare for the elderly: The effects of SSI on preretirement labor supply. *J. Public Econom.* 78(1–2):51–80.
- Perez-Truglia R, Troiano U (2018) Shaming tax delinquents. *J. Public Econom.* 167:120–137.
- Pomeranz D (2015) No taxation without information: Deterrence and self-enforcement in the value added tax. *Amer. Econom. Rev.* 105(8):2539–2569.
- Schultz PW, Nolan JM, Cialdini RB, Goldstein NJ, Griskevicius V (2007) The constructive, destructive, and reconstructive power of social norms. *Psych. Sci.* 18:429–434.
- Shang J, Croson R (2009) A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *Econom. J. (London)* 119:1422–1439.
- Slemrod J (2019) Tax compliance and enforcement. *J. Econom. Literature* 57(4):904–954.
- Social Security Administration (2015) Report on Supplemental Security Income non-medical redeterminations, fiscal year 2011. Social Security Administration, Baltimore.
- Social Security Administration (2019) Agency annual overpayment waivers report, fiscal year 2019. Social Security Administration, Baltimore.
- Social Security Administration (2020a) Agency financial report, fiscal year 2020. Social Security Administration, Baltimore.
- Social Security Administration (2020b) SSI annual statistical report, 2019. SSA Publication No. 13-11827. Social Security Administration, Baltimore.
- Thaler RH, Sunstein CR (2008) *Nudge: Improving Decisions About Health, Wealth and Happiness* (Yale University Press, New Haven, CT).