Journal of Public Economics 245 (2025) 105370

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Can changes in disability insurance work incentives influence beneficiary employment? Evidence from the promoting opportunity demonstration☆

Michael Levere [a,*], David Wittenburg [b], John T. Jones [c]

[a] Colgate University, United States
[b] Westat, United States
[c] Social Security Administration, United States

A B S T R A C T

We study how disability beneficiary work behavior responds to a rule change that replaces a cash cliff—a threshold above which benefits reduce to zero—with a benefit offset ramp—where benefits are gradually phased out. Using a randomized controlled trial with over 10,000 Social Security Disability Insurance beneficiaries who voluntarily enrolled in the demonstration, we find precisely estimated null effects on earnings, income, and benefit amounts. An analysis of mechanisms indicates that administrative burden, the limited size of the incentive, and individual and systemic barriers to employment for people with disabilities likely contributed to the limited impacts.

1. Introduction

There is strong policy interest in assessing whether Social Security Disability Insurance (SSDI) program rules affect beneficiary work activity. Extensive research has shown that some SSDI beneficiaries have lingering work capacity, despite the definition of disability requiring them to have an inability to "engage in substantial gainful activity" (Chen and Van der Klaauw, 2008; Maestas et al. 2013; French and Song 2014; Gelber et al. 2017). With concern that current SSDI rules linking earnings and benefits—including a cash cliff that leads beneficiaries to lose all benefits if earnings are sustained above a key threshold for an extended period[1]—might disincentivize work, policymakers have long considered whether offering greater financial incentives for people to work might in fact promote returns to work. For example, discussions of a $1 for $2 benefit offset to replace the cash cliff date back 25 years in the United States (Kostøl and Mogstad, 2014).

We explore the impact of modifying work incentive rules for SSDI beneficiaries by replacing the cash cliff with a benefit offset from the Promoting Opportunity Demonstration (POD). This "benefit offset" reduces benefits by $1 for every $2 in earnings above a certain amount. In addition, the rule changes also simplified other provisions of current rules to allow the benefit offset to apply immediately rather than have an extended period for beneficiaries to test work activity before work affects benefits. SSA applied the benefit offset monthly, thus requiring beneficiaries to report their earnings each month.[2] Beneficiaries with earnings sufficiently high to use the offset[3] might work more when the uncertainty of a cash cliff is removed, but the actual incentives to work vary depending on a person's earnings prior to entering POD. For some

☆ The research reported herein was performed pursuant to contract SS00-17-60008 from the U.S. Social Security Administration (SSA). The findings and conclusions are solely those of the author(s) and do not represent the opinions or policy of SSA or any agency of the Federal Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation, or favoring by the United States Government or any agency thereof.

* Corresponding author.
E-mail address: mlevere@colgate.edu (M. Levere).

[1] Despite a seemingly strong incentive to avoid work above this threshold, there is limited evidence in the United States that this cash cliff provision substantially affects work activity (Schimmel et al. 2011; Weathers and Hemmeter 2011; Gubits et al. 2018). Outside the United States, evidence from Austria indicates people bunch earnings just below another key earnings threshold (Ruh and Staubli 2019).

[2] Beneficiaries with the same earnings each month did not need to report their earnings every month, as the offset adjustment was based on the most recent amount of earnings reported. However, beneficiaries were still encouraged to report earnings each month, particularly because earnings of people with disabilities can be particularly volatile (Jolly and Wagner 2023).

[3] Throughout the rest of the paper, we refer to beneficiaries with earnings above the offset threshold as using the offset to simplify the language.

https://doi.org/10.1016/j.jpubeco.2025.105370
Received 5 April 2024; Received in revised form 27 March 2025; Accepted 1 April 2025

with high earnings whose benefits were already suspended, this offset might be predicted to reduce work, whereas for others the predicted impact on work is either ambiguous or no effect. In our analysis, we estimate subgroup impacts corresponding to these various levels of earnings. Regardless of prior earnings, however, to maximize income under these new offset rules, beneficiaries must both understand the rules and sustain their work effort.

We use a randomized controlled trial with more than 10,000 self-selected SSDI beneficiaries to explore the impact of the POD rule changes two years after enrollment. Treatment group members were subject to the benefit offset, while control group members faced current rules (including the cash cliff).[4] Because of the randomized design, any differences in outcomes between the treatment and control groups represent the causal effect of the rule changes. We examine impacts on outcomes related to benefit amounts, employment, and earnings from SSA administrative data. We use POD survey data to capture information not available in the administrative records to provide additional information on impacts related to employment activity (e.g., work search), income, health, and receipt of other program benefits. We also draw on data collected through the POD implementation in supplemental analyses to help understand the mechanism behind the impacts (or lack thereof).

These rule changes had limited impacts over the two years after enrollment. Average earnings, SSDI benefits, and income are essentially identical between the treatment and control groups. The estimated impacts are also sufficiently precise to rule out substantive changes. We did find one statistically significant impact on the percentage of people who had "substantive earnings" (measured as earning above roughly $15,000). In this case, treatment group members were about 1 percentage point more likely, or 10 percent relative to the control group mean, to have substantive earnings. However, once we correct for multiple hypothesis testing by computing sharpened False Discovery Rate q-values (Anderson 2008), this impact is no longer significant. Impacts on other outcomes related to employment activity, health, and receipt of other program benefits were also mostly limited.

We conducted subgroup analyses to assess if impacts differed based on prior earnings level, where the theoretically predicted change in work activity differs depending on one's prior earnings. Yet we mostly found no effects on work, except for a 4.4 percentage point increase in the rate of substantive earnings for those with some earnings (representing a 38 percent increase relative to the control group). This finding is notable because the simple theory would have predicted no change in employment activity for this group. In contrast, even for a group where theory would predict lower work activity, we found no impact on employment outcomes.

We then explore the reasons why impacts on earnings and benefit amounts were limited despite the substantial change in rules, highlighting three potential mechanisms that contributed to limited impacts. The first mechanism is administrative burden, which can be any sort of hassle or challenge people face when interacting with a government program (Herd and Moynihan 2019). These can include learning, compliance, and psychological costs. Though administrative burden decreased because of the rule simplifications in the demonstration, significant learning costs remained. Only 46 percent of treatment group members correctly understood that the benefit offset reduced their monthly benefits for earnings above a threshold, the fundamental premise of the demonstration. Additionally, overpayments—which require the beneficiary to owe back benefits received during a period where benefits should instead have been reduced because of work activity—affected over one in five treatment group members, or nearly three in four among those with benefit adjustments.[5] Overpayments may especially exacerbate challenges with understanding how work activity affects benefits. Finally, the short two-year timeframe may have been too short for treatment group members to fully adapt to the new rules or for work incentives to influence employment intentions. Though the demonstration clearly reduced burdens for beneficiaries along multiple dimensions, these factors indicate that enough burdens may have remained such that it limited impacts.

A second issue was the size of the incentive itself. For some, the new offset rules would theoretically be predicted to reduce work activity. People whose benefits were suspended at the time of enrollment would likely reduce their labor supply both because of an income effect (additional income from benefit payments) and a substitution effect (each $1 of earnings now leads to only $0.50 of additional take-home pay). Another group—those who had earnings such that their benefit would be reduced under the new rules but whose benefit would have been unaffected under current rules—has ambiguous predicted labor supply effects (the substitution effect predicts a reduction in labor supply while the income effect predicts an increase in labor supply). Empirically, we found no effects for either of these subgroups (though caution that both subgroups are quite small).

A third contributing factor was systemic and individual barriers to employment. At enrollment, 90 percent agreed it was difficult to work because of a physical or mental condition. Employment activity was limited among control group members: more than half of people had no earnings over the two-year study period, while only 10 percent of people had earnings above about $15,000 per year. Together with evidence around limited work activity for people with disabilities generally (e.g., Livermore and Honeycutt, 2015), potentially because of discrimination (e.g., Bellemare et al. 2023), this suggests that even if people wanted to work to take advantage of the rules, they may not have been able to.

Our paper adds to existing literature on the complexities of enhancing work activity among SSDI beneficiaries. Studies like Maestas et al. (2013), French and Song (2014), Gelber et al. (2017), and Moore (2015) indicate limited work activity among these beneficiaries, attributed partly to SSDI participation itself. These studies exploit exogenous factors related to program administration (such as variation in judge stringency) to understand the casual effects of SSDI receipt. Nichols et al. (2021) further elucidate this by discussing the modest impact of SSA's work incentives, potentially due to SSDI beneficiaries' limited work capacity. Our analysis aligns with these findings, underscoring the restricted work activity in this group.

Our findings also provide insights into the importance of administrative burden in government benefits programs. Typically, the literature highlights ways that administrative burden affects program enrollment. For example, Deshpande and Li (2019) show how field office closures, which make it harder for people to apply for benefits, reduce applications to Supplemental Security Income (SSI). Our findings contribute to a relatively new literature on redemption costs (Barnes 2021), which highlight ways that administrative burden can prevent people from effectively using a program.

[4] The treatment group technically consisted of two separate treatment groups. One treatment group faced termination if they had 12 consecutive months of earnings sufficiently high to lead benefits to fall to zero. The other treatment group did not face termination due to excess earnings. However, work activity and use of the benefit offset were nearly identical across the two treatment groups. Further, fewer than 1 percent of the termination-possible treatment group faced the termination provision. Therefore, we pool together these two treatment groups throughout this paper and conduct analysis as if there was only a single treatment group.

[5] Importantly, the overpayments for those subject to the new rules represent a very different type of burden than these beneficiaries likely otherwise would have experienced under current SSDI rules. Under current rules, overpayments are less frequent, but when they occur are for large sums of money (Hoffman et al. 2019). Under the demonstration rules, overpayments were much more frequent but for a much smaller dollar amount.

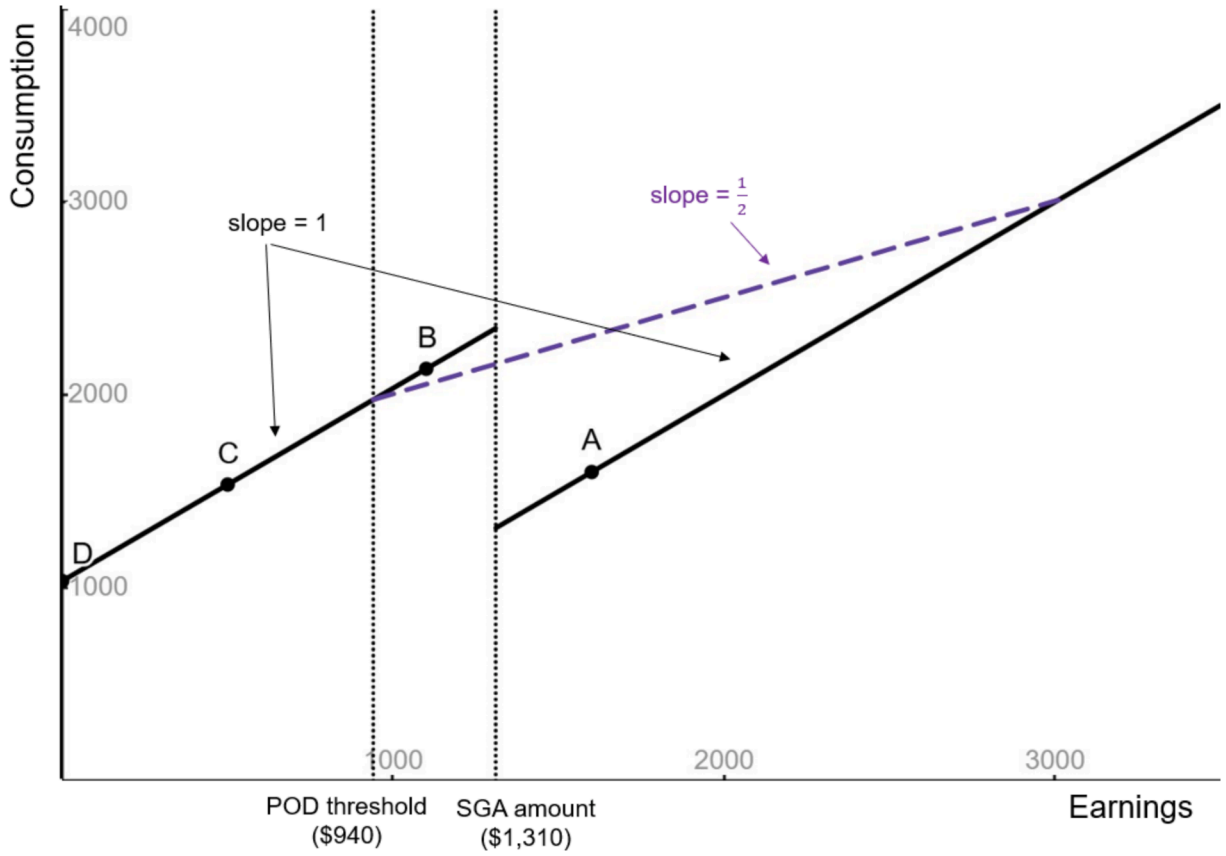

Fig. 1. Relationship between earnings and consumption under current rules and POD rules Note: The black solid line shows the relationship between earnings and consumption for an SSDI recipient under current rules with a monthly benefit amount of $1,035 (the average for those in the sample). With no earnings, their income is captured entirely by their SSDI benefit payment. For each additional dollar of earnings, consumption goes up by $1 as the benefit amount remains constant, as long as their earnings are below the Substantial Gainful Activity amount. At that point, their SSDI benefit payment falls to $0, leaving them with only labor earnings. For each additional dollar of earnings from there, consumption again goes up by $1 (for simplicity, ignoring the potential for any income taxes). The purple dashed line shows the change in this relationship under the benefit offset: starting at a lower threshold of the Trial Work Period amount, each $1 of earnings only leads to a $0.50 increase in consumption as the benefit amount is phased out at a rate of $1 for every $2 of earnings. Once the SSDI benefit is fully phased out (at a point equal to the POD threshold of $940 plus double the benefit amount, or $3,010 here assuming a $1,035 benefit amount), the beneficiary resumes seeing consumption increase by $1 for every additional $1 of earnings. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Institutional context

A. Overview of current SSDI rules

To qualify for SSDI benefits, an individual must be unable to engage in work that constitutes substantial gainful activity (SGA), defined as $1,310 per month in 2021. Earnings above the SGA amount are typically considered evidence that the beneficiary does not have a work-limiting impairment and therefore is ineligible to receive SSDI benefits.

In any 60-month period, beneficiaries are allowed to have earnings in excess of the SGA amount for up to 9 months without consequence. These 9 months are called the Trial Work Period (TWP). After completing the TWP, a beneficiary immediately enters a period called the Extended Period of Eligibility (EPE). The EPE is a 36-month re-entitlement period, during which benefits are suspended for months in which earnings exceed the SGA amount and full benefits are paid for months in which earnings fall below the SGA amount (i.e., they are not adjusted for any earnings below that amount).[6] The one exception is a three month grace period, which refers to the first month after the start of the EPE in which a beneficiary earns above the SGA amount and the two subsequent months. This grace period often starts immediately after completing the TWP, such as if the earnings that led to the completion of the TWP continue and were at a level above the SGA amount.

After the EPE ends, SSA terminates benefits if earnings exceed the SGA amount (unless the beneficiary has not yet used the three month grace period, in which case the grace period could apply). Otherwise, benefit payments continue in full. If benefits are terminated due to SGA, beneficiaries can seek expedited reinstatement of benefits at any point during the 60 months following their notification of benefit termination by SSA. Substantive earnings activity among SSDI beneficiaries can also affect their Medicare eligibility after an extended period.[7]

SSA has a long-standing interest in understanding options to promote employment among SSDI beneficiaries. For example, SSA tested a different benefit offset that also eliminated the cash cliff as part of the Benefit Offset National Demonstration (BOND). The changes in program rules under BOND applied at a higher level of earnings and only after the beneficiary had completed the TWP and the grace period. BOND had minimal effects on earnings and led to increases in SSDI benefit

[6] In making this SGA determination, SSA uses an adjusted measure of earnings that deducts SSA-approved impairment related work expenses and other noncountable income such as sick pay, vacation pay, and subsidies.

[7] SSDI beneficiaries become eligible for Medicare Part A benefits (and can pay a monthly premium to receive Medicare Part B benefits) 24 months after SSDI eligibility. Beneficiaries with cash benefits terminated based on the performance of SGA generally lose their Medicare benefits 93 months after completion of the TWP.

payments (Gubits et al. 2018). The design of POD, which was Congressionally mandated, drew on lessons from that previous experience. For example, POD addresses the perception that BOND rules were complex by using a simplified set of administrative adjustments in implementing the offset. This attempt is important because it could reduce the administrative burden to treatment group members in understanding the implications of work rules. However, POD also included other provisions that could increase the administrative burden to treatment group members, especially monthly earnings reporting (which represent a type of compliance cost).

B. The Promoting Opportunity Demonstration: theoretical considerations

The rules for POD included a benefit offset ramp and modified other current rule provisions. This offset reduced benefits by $1 for every $2 of monthly earnings higher than the POD threshold (equal to the TWP amount, which was $940 in 2021).[8] The benefit offset applied immediately to monthly earnings, eliminating the TWP, grace period, and EPE. This change made it more straightforward to describe how earnings affected benefit adjustments, and thus likely reduces administrative burden. However, it also required that treatment group members report earnings monthly for their benefits to be adjusted promptly, which introduced a different administrative burden.[9]

To explore the theoretical predictions of the offset on work behavior, we draw on the neoclassical model of labor supply. Fig. 1 shows the relationship between earnings and consumption for an SSDI beneficiary (for simplicity, ignoring the TWP and grace period and any income taxes), showing earnings on the x-axis and consumption on the y-axis. The assumed SSDI benefit amount in the figure is $1,035, the average SSDI benefit payment among all enrollees (Table 1). The black line shows this relationship for such a beneficiary under current SSDI rules. For someone who does not work at all, their consumption comes solely from their SSDI benefit. As earnings increase, consumption also increases by the same amount given that the SSDI benefit amount is fixed. However, if their earnings exceed the SGA amount, their SSDI benefit is fully lost (represented by the fall in consumption at the SGA amount of $1,310). After this point, each additional $1 of earnings also increases consumption by $1. Under POD, the benefit offset takes effect at the TWP amount of $940 (below than the SGA amount). At this point, each $1 of earnings only increases consumption by $0.50 because of the benefit offset (represented by the dashed purple line). After the SSDI benefit fully phases out to $0—which occurs at double the benefit amount plus the POD threshold amount of $940 (in the figure with a $1,035 benefit, this corresponds to $3,010)—each additional $1 of earnings again increases consumption by $1.

Under this simple model, the theoretical effects of the offset on work behavior vary depending on a beneficiary's earnings because of differing income and substitution effects. A beneficiary at point A,[10] who would have had benefits suspended under current rules because of work exceeding the SGA amount, would have less incentive to work. The reintroduced benefit would reduce labor supply through an income effect, while the phase-out and 50 percent marginal tax rate would also reduce labor supply through a substitution effect. The theoretical

Table 1
Balance across treatment and control groups at baseline.

Variable	Treatment group	Control group	*p*-value
Basic characteristics			
Female	55.2	54.5	0.473
Mean age (years)	47.4	47.4	0.794
Preferred language is Spanish	2.4	3.1	0.053
Primary diagnosis			
Neoplasms	2.9	2.9	0.951
Mental disorders	37.9	39.0	0.280
Back or musculoskeletal system	20.1	20.4	0.673
Nervous system disorders	6.5	6.1	0.431
Circulatory system disorders	5.7	6.0	0.553
Genitourinary system disorders	4.3	4.1	0.655
Injuries	3.9	3.8	0.896
Respiratory	1.7	1.7	0.780
Several visual impairments	2.3	2.3	0.848
Digestive system	1.4	1.5	0.632
Other impairments	10.7	9.5	0.049
Beneficiary program characteristics			
Mean SSDI duration (months)	113.3	115.5	0.174
Monthly SSDI benefits ($)	1035.6	1032.9	0.756
Has representative payee	6.6	7.5	0.106
Concurrent SSI receipt	18.4	17.8	0.442
Employment history			
History of TWP-level earnings since 2014	19.0	19.5	0.496
History of SGA-level earnings since 2014	14.7	15.3	0.380
Had a Ticket assigned in last four years	12.8	12.0	0.242
Work status at baseline			
Currently employed	24.0	25.1	0.158
Seeking work	23.9	23.5	0.653
Neither employed nor seeking work	52.1	51.4	0.469
Monthly earnings over $1,000	13.1	13.0	0.976
Expects to work in the next year [a]	61.3	61.0	0.763
Received job training, job coaching, or support services	15.9	17.4	0.043
Agrees with statement:			
Difficult to work because fear losing disability cash benefits	57.7	57.4	0.790
Difficult to work because of a physical or mental condition	89.5	88.2	0.044
Difficult to work because of unreliable transportation	34.9	33.6	0.196
Difficult to work because caring for children	15.7	16.4	0.415
Difficult to work because don't have needed skills or training	31.8	32.2	0.696
Family income category:			
Less than $10,000	31.7	32.8	0.262
$10,000 to less than $20,000	36.7	33.8	0.004
$20,000 to less than $30,000	12.6	13.9	0.072
$30,000 to less than $50,000	10.3	10.5	0.679
$50,000 or more	8.8	9.0	0.716
Observations	6,700	3,370	10,070

Source: Authors' calculations using Mathematica's POD recruitment and enrollment system, SSA program records, POD Implementation Data System, and the POD baseline survey.
Note: Unless otherwise noted, all table entries are percentages. The *p*-values in the final column of the table are based on joint tests for differences between the T1, T2, and C groups. These tests compare means for continuous variables and proportions for binary variables.

[8] POD also includes special provisions for beneficiaries who have Impairment-Related Work Expenses (IRWE). For such beneficiaries, the threshold is defined as the greater of the TWP amount and a beneficiary's IRWE (up to a maximum of the SGA amount). However, such beneficiaries are rare, with fewer than one percent of POD participants having a threshold higher than the TWP amount.

[9] Importantly, though, under current rules beneficiaries need to report whenever they start or stop work, if duties, hours, or pay changes, or there are impairment-related work expenses. Thus, this may not represent a substantial change to the reporting process for beneficiaries who work.

[10] In the figure, point A is shown at earnings of $1,600. Under current rules, consumption would be $1,600 because their SSDI benefit amount would fall to $0. Under POD rules, consumption would be $2,305 because their SSDI benefit amount would be $705 (a $330 reduction in benefits from their benefit amount of $1,035, given earnings exceeding the threshold of $940 by $660). However, the general theory applies to anyone with earnings higher than the SGA amount of $1,310.

prediction is ambiguous for a beneficiary at point B,[11] who would have had a full benefit check under current rules but a reduced benefit amount under POD. The reduction in the benefit amount increases labor supply through an income effect, while the 50 percent marginal tax rate reduces labor supply through a substitution effect. Beneficiaries who have a modest amount of work (point C)[12] or no work at all (point D)—representing almost 90 percent of the people enrolled in the demonstration—would not be predicted to change their work behavior given that the benefit offset does not affect the relationship between work and income at their current level of work activity.

However, this simplicity of this model has conceptual shortcomings which make predicting the theoretical effects of the benefit offset challenging. Most importantly, the model assumes that an employee can directly control their hours and work precisely the number of hours that they choose to. In reality, a beneficiary may face lumpy work hours – for example, a choice may be to work exactly a forty-hour work week or not work at all. In this instance, the cash cliff may have been a powerful motivator that led someone to choose to not work under current rules (Ruh and Staubli (2019) show that people manipulate their earnings to avoid discrete drops in their disability benefits, though evidence from Schimmel et al. (2011) and Gubits et al. (2018) suggest that bunching earnings just below the SGA amount is rare in the US context). It also assumes that people accurately understand the rules and the precise way that work activity will influence benefit amounts and thus total consumption.

The removal of that cash cliff may therefore incentivize more work activity even among those who, in the simple model of Fig. 1, may not have been predicted to change their labor supply. The removal of the cash cliff may also lead to greater labor supply even among people working a small amount—despite no predicted change under the simple model—for two reasons. First, such people may misunderstand the amount of work that corresponds to the cash cliff. Indeed, in Section 6.A, we show that beneficiary understanding of the rules is extremely limited. Second, such people may be risk averse and fear that additional work may lead to the loss of benefits, whereas the removal of the cash cliff may assuage those fears. At baseline, about 60 percent of people said they agreed that it was difficult to work because of a fear of losing disability cash benefits. The share agreeing with that statement was stable regardless of prior work history. The existence of the TWP and the grace period further complicate the picture by removing the cash cliff for several months.

On net, POD therefore might have ambiguous impacts on benefits and earnings, particularly depending on a beneficiary's recent connection to work. We estimate subgroup analyses by a beneficiary's baseline level of earnings to see if the intervention has heterogeneous effects. Specifically, we group people by whether they fell into categories associated with points A, B, C, and D in Fig. 1: benefits suspended at some point in the year before enrollment (point A), not suspended but earned above the threshold for the benefit offset in the year before enrollment (point B), some earnings less than the threshold for the benefit offset in the year before enrollment (point C), no earnings in the year before enrollment (point D). The predicted effects of the benefit offset at point A are unambiguous – given a person already worked enough to give up their SSDI benefit, the income and substitution effects both theoretically reduce labor supply. However, the predicted effects at points B, C, and D are more ambiguous—both because of opposite signed income and substitution effects and because the cash cliff under current rules may have impacted their baseline work activity in ways that make it difficult to draw inference from the static model.

C. The Promoting Opportunity Demonstration: experimental implementation

SSA created a separate infrastructure to implement this benefit offset and support beneficiaries through the process. SSA funded an implementation team that created a system to collect earnings information (e.g., pay stubs) from treatment group members. This team then sent the information to SSA, who recorded those earnings as received and, in most cases, used an automated system to process benefit adjustments. At the end of the year, SSA checked if an adjustment was necessary for earnings not reported.[13]

Recruitment efforts included both direct and indirect outreach to inform beneficiaries of the rule changes. Direct outreach included informational mailings sent to potentially eligible beneficiaries, while indirect outreach included efforts to raise awareness through a website and communications with community organizations that serve SSDI beneficiaries. Benefits counselors also offered support throughout the demonstration. Counselors contacted treatment group members upon enrollment to provide information about the new POD rules. The counselors also provided referrals and more in-depth work incentive counseling throughout the demonstration based on the treatment group members' needs; the latter type of counseling is also available under current program rules.

10,070 working-age SSDI beneficiaries in the eight POD states (Alabama, California, Connecticut, Maryland, Michigan, Nebraska, Texas, and Vermont) volunteered for POD and provided written informed consent between January 2018 and January 2019. This represents 2.4 percent of those solicited to sign up for the demonstration. Once enrolled, people were eligible to use the benefit offset until June 2021, when the demonstration ended, at which point they returned to current-law rules. A key feature of the demonstration authority was that participation was voluntary, and, hence, participants could withdraw from the demonstration at any time.[14]

The COVID-19 pandemic overlapped with part of the demonstration period, during which SSA made several adjustments for SSDI beneficiaries generally and POD treatment group members specifically to maintain continuity of services. These changes allowed the demonstration (and current SSDI services) to continue despite broad disruptions from the pandemic. First, SSA severely limited in-person services at its local field offices and reprioritized workloads at the start of the pandemic in March 2020. Second, SSA added protections to help avoid disruptions to benefit payments for all SSDI beneficiaries. Finally,

[11] In the figure, point B is shown at earnings of $1,100. Under current rules, consumption would be $2,135 because they would receive their full SSDI benefit amount. Under POD rules, consumption would be $2,055 because their SSDI benefit would fall by $80 to $955 given their earnings exceed the threshold of $940 by $160. However, the general theory applies to anyone with earnings higher than the POD threshold of $940 and below the SGA amount of $1,310.

[12] In the figure, point C is shown at earnings of $500. Under both current rules and POD rules, consumption would be $1,535 because they would receive their full SSDI benefit amount regardless of the rules. The general theory applies to anyone with positive earnings that are less than the POD threshold of $940.

[13] Specifically, SSA used the POD automated data system to sum each treatment group member's monthly earnings reports submitted across all months in the year and compared them with the total annual gross earnings from Internal Revenue Service records. This process allowed SSA to determine the SSDI benefits that should have been paid to each POD treatment group member during the previous calendar year and compare it to the actual amount of SSDI benefits paid.

[14] About 8 percent of those in the treatment group withdrew from the demonstration. Most people who withdrew did so shortly after enrolling. Of those who withdrew, about 35 percent cited a lack of interest in the work incentives, such as POD not being beneficial because their earnings were in the range where total income would be lower under the POD rules than current rules. Another 25 percent cited being unlikely to work as a reason to withdraw, such as being too disabled to work or not interested in working. Our analysis of administrative data includes those who withdrew based on their initially assigned treatment arm, and thus represents an intent-to-treat analysis.

counselors called all treatment group members in the early phase of the pandemic to offer support, connect them to area resources, and inquire about changes in their employment status. Importantly, we do not find substantive differences in impacts over the pandemic period.[15]

Beneficiaries who enrolled in the demonstration differed from the general SSDI beneficiary population, particularly in their interest in work (Appendix Table 1). For example, 15 percent of POD enrollees had earnings at or above the SGA amount in any month since 2014, which was about 2.5 times the rate for non-volunteers. This group, which would be approaching benefit suspense depending on their completion of the TWP, is particularly notable given that POD would be predicted to reduce work activity for anyone who has benefits already in suspense at enrollment. POD enrollees also differed from beneficiaries who did not volunteer along other characteristics, though many of these differences may also stem from enrollees' stronger connection to work. For example, POD enrollees were younger and were more likely to have a mental disorder than non-volunteers (von Wachter, Song, and Manchester 2011 show that those with mental disorders have greater labor force attachment). The implication of these differences is that the evaluation findings are specific to the sample of POD enrollees. In other words, our findings represent the impacts on the beneficiary population who decided to enroll and do not necessarily generalize to a nationally representative population of SSDI beneficiaries. Nonetheless, because those who enrolled have a strong interest in working, we may have been more likely to find a significant impact of the demonstration than among a randomly sampled group of SSDI beneficiaries. In considering the findings, it is important to reiterate the low overall rate of sign-up for the demonstration, which was only 2.4 percent. However, this low rate is comparable to participation in other SSA demonstrations, like BOND (which enrolled 5.4 percent of those contacted).

3. Data

Our data included information from SSA program records, earnings reported to the IRS, the POD Implementation Data System, and three surveys. Together, these data enabled us to examine offset use, assess understanding of earnings rules, assess experiences with improper payments, and to estimate program impacts on employment, benefits, and other outcomes.

Data on SSA program participation and earnings comes from several SSA data sources, which we used to construct all the primary outcomes for the impact analysis. We use the Master Beneficiary Record and the Supplemental Security Record to track monthly SSDI and Supplemental Security Income (SSI) program participation, respectively. We construct outcomes corresponding to benefit amounts due in the first 24 months after enrolling in the demonstration.[16] Our employment and earnings measures from the Master Earnings File represented average annual earnings reported to the IRS. The annual earnings data covered 2019 and 2020, which encompassed the two calendar years after the year of enrollment.[17][18] We also construct measures of average annual income, which include earnings plus SSDI and SSI benefit amounts due in 2019 and 2020.[19] Our analysis of improper payments is based on monthly snapshots from the Master Beneficiary Record. To identify improper payments, we look at whether the cash benefit due at the time the payment was made is different from the cash benefit due for that month as of June 2021.[20]

We also use information from the POD Implementation Data System and three surveys to assess additional outcomes and probe the mechanisms driving our findings. Data on benefit offset usage comes from the Implementation Data System,[21] which we report on for the first 24 months after enrolling in the demonstration. Our survey data efforts included a baseline survey and two follow-up surveys. Beneficiaries had to complete the baseline survey to enroll in POD. The two follow-up surveys included content on follow-up activities one and two years after random assignment. Both surveys captured information about enrollees' employment, understanding of program rules, attitudes about work, income, health and functional status, and health insurance. More than 80 percent of surveyed beneficiaries completed the follow-up surveys (84 percent for the first follow-up survey and 83 percent for the second follow-up survey). The first follow-up survey included a random sample of half of POD enrollees,[22] while the second follow-up survey included the full sample of POD enrollees. Response rates did not differ for the treatment and control groups.

4. Methods

We used a randomized controlled trial to test the impact of the modified rules. The 10,070 beneficiaries who enrolled in the demonstration were randomly assigned to either a treatment group (6,700 beneficiaries) or control group (3,370 beneficiaries). Random assignment means that control group members should represent a valid counterfactual for what outcomes of treatment group members might have looked like if not for their participation in POD.

We did not find any substantive differences across any major demographic, impairment or work history categories between control and treatment groups (Table 1).

Equation (1) shows our primary estimating equation. We estimated

[15] Specifically, our findings are similar for 2019 alone, 2020 alone, or the average across both years (the latter corresponding to the primary findings presented in this paper). Average earnings in 2020 were roughly similar to earnings in 2019 for both the treatment and control groups despite the pandemic.

[16] We explore SSDI benefit payments in particular as an outcome to better understand the extent to which reductions in benefit amounts for those who previously had been receiving a full benefit check are outweighed by increases in benefit amounts for those who previously might have had benefits suspended but now have their benefit reinstated.

[17] About 2 percent of participants were enrolled and randomly assigned in January 2019 (Hock et al. 2020). However, because these beneficiaries had to submit their enrollment materials before December 31, 2018, outcomes measured in calendar years 2019 and 2020 are still a good proxy for their experience in the first two years after enrollment. To maintain consistency, we essentially treated December 2018 as the month of enrollment for beneficiaries who enrolled in January 2019.

[18] Though benefit reductions under both POD and current rules are based on a measure of monthly earnings, that measure is not accessible for the purposes of research. Instead, the only data available to analyze impacts for the purposes of this paper is the annual data. Monthly earnings data would be helpful in assessing various aspects of people's responses to the benefit offset, though unfortunately such an analysis is not possible with the currently available data.

[19] This is of course only a partial measure of income, but captures what we are reliably able to measure from SSA administrative records.

[20] For more details on how we measure improper payments, see Appendix E and Section VI.3 in Wittenburg et al. (2022).

[21] The POD implementation data system, which the implementation team maintained, included information on POD related services, such as the provision of work incentive counseling, collection and submission of earnings information to SSA, offset use, and transition back to program rules at the end of subjects' POD participation period.

[22] We designed the random sampling procedure to guarantee that the characteristics of those who were selected to participate in the survey closely resembled the characteristics of all POD enrollees. The random selection plus similar characteristics of the survey sample means that the estimates from the survey data are representative of all POD enrollees.

impacts using an ordinary least-squares model with heteroskedasticity-robust standard errors.

$$y_i = \alpha + \beta T_i + \delta X_i + \varepsilon_i \quad (1)$$

The coefficient of interest is β, which measures the adjusted difference in means between the treatment and control groups in an intent-to-treat setting, only considering the group someone was initially assigned to, not whether they engaged with the intervention in any way or even withdrew from the demonstration. Because of the randomized design, β therefore represents the causal impact of POD. We control for several individual characteristics in X_i, including those used to stratify random assignment[23] and an array of baseline characteristics, such as sex (collected in administrative records), race (collected in the baseline survey), recent work activity and self-reported family income (also in the baseline survey), and three characteristics that exhibited chance statistical differences between the groups at baseline.[24] Our results are robust to the exclusion of these individual characteristics. Except for the characteristics used to stratify random assignment, including these characteristics as controls should only improve the precision of the impact estimates but they are not needed to generate unbiased results.

Our analysis includes administrative data for all POD enrollees and, for survey outcomes, the subset of enrollees who responded to surveys. For the latter, we use weights to produce estimates that reflect the impact of POD rules on all POD enrollees. Analysis weights for outcomes from the one-year follow-up survey account for survey sampling and nonresponse,[25] while the analysis weights for the two-year follow-up survey only account for nonresponse as the survey included all enrollees.

We mostly focus on four primary outcomes as the main assessment of POD's efficacy. We pre-specified these four primary outcomes mainly to avoid problems related to multiple comparisons. However, we also use the approach described by Anderson (2008) to compute a sharpened False Discovery Rate (FDR) q-value that directly adjusts for multiple hypothesis testing. We calculate these sharpened FDR q-values among the primary outcomes alone, the primary and secondary outcomes together, and for all primary outcomes across subgroups. This approach is consistent with research like Guess et al. (2023), which uses the same approach to correct for multiple hypothesis testing.

The four primary outcomes include continuous measures of earnings, benefit payments, and income, as well as an indicator for having substantive earnings, or about $15,000 in earnings (more precisely, the annualized SGA amount). Secondary outcomes include several employment-related outcomes and several health and health insurance related outcomes. For employment-related outcomes, we report on information collected from the survey (any employment in the past year, being employed or actively searching for a job, hours worked, fringe benefits), SSA program records (whether earnings were greater than several notable thresholds, such as $0), and vocational rehabilitation (VR) program records (whether people applied for or received VR services). We also analyze measures of physical and mental health based on the 12-item Short Form Survey developed from the Medical Outcomes Study (Hays et al. 1995), as well as whether the beneficiary reports having any health insurance coverage.

We also conducted subgroup impact analyses to assess whether the modified rules had differential impacts depending on a beneficiary's recent level of earnings.[26] We divided the enrolled beneficiaries into four groups based on their recent earnings history, each corresponding to one of the four points shown in Fig. 1 (as discussed in Section 2.B). We estimate the impact of assignment to the treatment group for each of the four subgroups, as well as run a statistical test for whether the impacts are different from each other.

5. Results

A. Primary impact results

The intervention had no overall impact on earnings (Table 2). Average annual earnings in 2019 and 2020 for treatment group members was $5,028, relative to $4,954 for the control group. The estimated difference of $74 represented about 1 percent of the control group mean, which implies no substantive change in outcomes. The 95 percent confidence interval covers the range of a decrease in earnings of $314 to an increase of $462. We can therefore rule out an effect on earnings of more than 10 percent.

Those assigned to the treatment group were 1 percentage point more likely to have substantive earnings, defined as annual earnings above approximately $15,000 (the annualized SGA amount), and this difference was statistically significant at the 10 percent level. We expected that those who would benefit most from the modified work rules would be those who could earn above this SGA amount on a continuous basis, the threshold at which the cash cliff applies. About 11 percent of the treatment group had such earnings, compared with 10 percent of the control group; the estimated difference of 1 percentage point represents a 10 percent increase relative to the control group mean. However, once correcting the p-value for multiple hypothesis testing, this finding is no longer significant: either when calculating a sharpened FDR q-value only among primary outcomes or when also including the secondary outcomes in the adjustment.

We also found no significant differences between the treatment and control groups in terms of SSDI benefit amounts or income. The average annual SSDI benefit amounts in the two years after enrolling for the treatment group was $11,874. Total income for treatment group members, which included total earnings plus SSDI benefit amounts and SSI payment amounts, was $16,768. For both measures, the control group mean was within 2 percent of the treatment group mean, which further underscores the interpretation of no impact. Standard errors are also precisely estimated: we can rule out changes of more than 4 percent in either direction.

The precisely estimated null impact on SSDI benefit amount is notable given that 30 percent of treatment group members used the benefit offset, as discussed further below. The rules associated with the benefit offset would lead some treatment group members to experience increased

[23] We stratified random assignment by state. Within each state, we also stratified in the following ways: first, if someone had one of three rare diagnoses (neoplasms, injuries, or severe visual impairments), we only stratified by the primary diagnosis. For everyone with a different diagnosis, we stratified by state, age groups (either ages 20 to 34, 35 to 44, or 45 and older), SSDI benefit duration (1 to 18 months, 19 to 36 months, and 36 months or more), and whether they reported earning over $1,000 per month in the baseline survey at enrollment.

[24] These three characteristics include: (1) the extent to which people agree that it is difficult to work because of a physical or mental condition; (2) whether people recently received job training, job coaching, or support services; and (3) whether the preferred language is Spanish. The first two are measured in survey data, while the third is measured in administrative data.

[25] The weights are the product of two terms: sampling weights and the survey nonresponse weights. The sampling weight (the first term) is determined by the probability of being sampled for that survey. Because we randomly sampled half the POD enrollees for the year-one follow-up survey, the sampling weight term in the overall weight is the same for all POD enrollees. To construct the survey nonresponse weight (the second term in the overall weight), we use a random forest algorithm. The algorithm uses observable baseline characteristics to predict the probability that each person responded to the survey. The nonresponse weight equals the inverse of the estimated response probability.

[26] In alternative analyses, we also consider subgroup analysis based on several other factors, including self-reported employment status at the time of enrollment, work expectation at the time of enrollment, level of education, age, primary impairment, or state of residence. However, we found essentially no notable differential impacts across subgroups for the primary outcomes, so to save space we do not report such findings in this paper. For the results of those analyses, please see Exhibits F.2 to F.6 in Wittenburg et al. (2022).

Table 2
Impacts on primary outcomes.

	(1) Average annual earnings	(2) Substantive earnings	(3) SSDI benefit amount	(4) Average annual income
Treatment	74	1.0*	135	220
	(199)	(0.6)	(105)	(196)
Control mean	4,954	10.0	11,739	16,548
Observations	10,070	10,070	10,070	10,070

Source: Authors' calculations using SSA program records.
Note: The number for treatment represents an estimate of β from equation (1), representing the estimated impact of being assigned to the treatment group. Standard errors, reported in parentheses, are robust to heteroscedasticity. All monetary values are in 2019 dollars. Substantive earnings is an indicator variable for whether the beneficiaries' average earnings over 2019 and 2020 exceeds the annualized SGA amount over the full two-year period. All outcomes are measured over the 2019 and 2020 calendar years. The exception is that SSDI benefit amounts are measured over the two years after POD enrollment.
***/**/* indicate estimate is significantly different from 0 at the 1/5/10 percent level using unadjusted p-values.
+++/++/+ indicate estimate is significantly different from 0 at the 1/5/10 percent level after adjusting for multiple hypothesis testing using sharpened false discovery rate q-values.

benefit amounts without any changes in earnings behavior (such as those who had completed the TWP and had substantive earnings). However, the new rules would also decrease the benefits of other treatment group members (for example, those in the TWP). Thus, a null effect may mask heterogeneity in that there were increases in benefit amounts for some and decreases for others. As shown next in Table 3, benefit amounts increased by nearly $3,000 per year for the small group of beneficiaries whose benefits were suspended at the time of enrollment.

We next analyzed heterogeneous impacts by beneficiaries' prior earnings level to correspond with the points shown in Fig. 1. As shown in Table 3, those with significant earnings (point A in Fig. 1) experience no change in employment outcomes though see a large increase in their SSDI benefit amount (which is differentially larger than the impact on SSDI benefits for the other subgroups). The latter finding is expected because many have their benefits reinstated with the offset ramp, rather than continuing to face a full loss of benefits with the cash cliff. However, the former is inconsistent with the theoretical prediction of reduced work activity both because of an income and substitution effect. For those with modest earnings (where the theory is ambiguous, corresponding to point B in Fig. 1), we find no significant impact on any of the primary outcomes.

In contrast, for those with some or no earnings (where the theory predicted no change), we see potentially positive effects on earnings and employment activity. Those with some prior earnings experienced a significant 4.4 percentage point increase in substantive earnings (earning above about $15,000 per year), which represented a 38 percent increase relative to the control group. This estimate is notable because it is both significantly different from zero and significantly different from the estimated impacts on substantive earnings for the other subgroups. It also continues to be significant even after adjusting for multiple comparisons. For those with no earnings, the point estimate on average annual earnings, while not significant, also represents a change of more than 10 percent.

B. Supplemental impact results

POD had relatively limited effects on most secondary outcomes (Table 4).[27] Treatment group members were about 3 percentage points more likely than control group members to report that they were either employed or actively searched for a job in the year before the survey. This impact represented a 5 percent increase relative to the control group mean. We also considered effects on various earnings levels throughout the income distribution using IRS data: anything above $0; above $1,000; above double the annualized SGA amount (i.e., double the value required for substantive earnings in Table 2); and above triple the annualized SGA amount. We found no significant impacts at these earnings thresholds. Point estimates were also close to zero at all earnings thresholds, especially when considered in percentage terms relative to the control group mean. We also found that treatment group members

Table 3
Impacts on primary outcome, by recent earnings history at enrollment.

	(1) Significant earnings	(2) Modest earnings	(3) Some earnings	(4) No earnings	(5) *p*-value for equality across groups
Panel A. Average annual earnings					
Treatment	783	−1,528	508	258	0.428
	(2,470)	(1,159)	(431)	(173)	
Control mean	23,381	14,071	6,753	2,090	
Panel B. Substantive earnings					
Treatment	−0.7	−0.9	4.4***++	0.6	0.078
	(5.5)	(3.2)	(1.4)	(0.5)	
Control mean	61.0	30.0	11.6	4.0	
Panel C. SSDI benefit amount					
Treatment	3,258***+++	255	−62	−47	0.002
	(865)	(436)	(229)	(117)	
Control mean	5,154	9,790	12,100	12,225	
Panel D. Average annual income					
Treatment	4,195**	−1,273	515	198	0.117
	(2,148)	(1,065)	(443)	(186)	
Control mean	28,272	23,627	18,550	14,245	
Observations	290	774	2,444	6,562	

Note: Divides all enrollees into four subgroups defined by earnings from 2017, the year before enrollment. Column (1) for significant earnings includes people whose SSDI benefit payments had recently been suspended. Column (2) for modest earnings includes people whose recent earnings had exceeded the Trial Work Period amount, but who did not have benefits suspended. Column (3) for some earnings includes people who had some recent earnings, though where such earnings did not exceed the Trial Work Period amount. Column (4) for no earnings includes people who had no earnings in the year before POD enrollment. The number for treatment represents an estimate of β from equation (1), representing the estimated impact of being assigned to the treatment group, for those within each subgroup. Standard errors, reported in parentheses, are robust to heteroscedasticity. The differential impact *p*-values in Column (5) come from a test of whether the impact estimate across all four subgroups is equal. All monetary values are in 2019 dollars. Substantive earnings is an indicator variable for whether the beneficiaries' average earnings over 2019 and 2020 exceeds the annualized SGA amount over the full two-year period. All outcomes are measured over the 2019 and 2020 calendar years. The exception is that SSDI benefit amounts are measured over the two years after POD enrollment.
***/**/* indicate estimate is significantly different from 0 at the 1/5/10 percent level using unadjusted p-values.
+++/++/+ indicate estimate is significantly different from 0 at the 1/5/10 percent level after adjusting for multiple hypothesis testing using sharpened false discovery rate q-values.

[27] We focus here on findings in the two-year follow-up survey, which included all enrollees. Results from the one-year follow-up survey are mostly similar, and are available upon request or in Wittenburg et al. (2022).

Table 4
Impacts on supplemental outcomes.

	Control group mean	Estimated impact	Standard error	Sample size
Employment-related outcomes				
Any employment in past year	33.5	0.9	1.0	7,842
Employed or actively searching for a job	50.9	2.5**	1.1	7,867
Any positive earnings (SSA program records)	45.0	1.1	0.9	10,070
Earnings above $1,000 (SSA program records)	37.8	0.4	0.9	10,070
Earnings above 2x annualized SGA amount (SSA program records)	4.2	−0.3	0.4	10,070
Earnings above 3x annualized SGA amount (SSA program records)	1.5	−0.0	0.2	10,070
Hours worked per week at most recent job	8.0	0.4	0.3	7,842
Any fringe benefits offered at most recent job	17.4	1.5*	0.8	7,842
Applied for VR services in first two years after enrollment (VR program records)	2.8	1.3***++	0.4	10,070
Received VR services in first two years after enrollment (VR program records)	4.0	0.7*	0.4	10,070
Health and health insurance related outcomes				
Physical health aggregate score	33.9	0.1	0.3	6,971
Mental health aggregate score	39.3	0.1	0.3	6,971
Has any health insurance coverage	98.5	0.2	0.3	7,732

Source: Authors' calculations using SSA program records.
Note: The estimated impact represents an estimate of β from equation (1). Unless otherwise noted, all table entries are percentages for means or percentage points for impact estimates. Unless otherwise noted, all data are from the year two follow-up survey. Data from the follow-up survey can be missing owing to item-level non-response. Data from the follow-up survey are weighted using survey non-response weights to account for the people who were sent the survey but did not complete it. The administrative data includes all initial participants in the demonstration. Standard errors, reported in parentheses, are robust to heteroscedasticity.
***/**/* indicate estimate is significantly different from 0 at the 1/5/10 percent level using unadjusted p-values.
+++/++/+ indicate estimate is significantly different from 0 at the 1/5/10 percent level after adjusting for multiple hypothesis testing using sharpened false discovery rate q-values.

had more active engagement with VR services than control group members. Those in the treatment group were 1.3 percentage points more likely to apply for and 0.7 percentage points more likely to use VR services. Though these differences are small in magnitude, they are large relative to the control group mean (about 50 and 20 percent, respectively). Notably, the impact on VR application is the only outcome that maintains significance in either Table 2 or Table 4 after adjusting for multiple comparisons. In contrast, we found no difference on most outcomes related to health and health insurance. For example, about 99 percent of both treatment and control group members had health insurance coverage, and average physical and mental health scores were similar.

6. Mechanisms

In this section, we explore potential reasons that might explain the limited impacts on primary outcomes presented in Section 5. We point to three main factors: (1) administrative burden (specifically, learning costs); (2) the incentive not being big enough; and (3) individual or systemic barriers to employment for people with disabilities more generally. We present evidence related to each of these three main barriers, suggesting that each likely play a role in the limited impacts. However, because the analyses in this section are primarily descriptive (i.e., they do not rely on the randomized design of the intervention), they should be viewed as suggestive in nature.

A. Administrative burden

The rule changes associated with the benefit offset simplified existing rules and thus limited learning costs and reduced administrative burden in several ways. For example, the rules eliminated the TWP and grace period associated with current SSDI rules, which a beneficiary would have to go through before facing the cash cliff.[28] Beneficiaries in the control group (for whom these rules continued to apply) had difficulty understanding these rules: only about one-quarter knew that a TWP existed in which benefits remain unchanged regardless of earnings (Table 5, Panel B). POD also eliminated work-related continuing disability reviews.

The rules also changed the burden associated with improper payments, which occur when SSA pays beneficiaries more (or less) in SSDI benefits than they were entitled to based on work activity. Under current rules, improper payments are quite rare, but overpayments impose a substantial financial burden to beneficiaries who experience them: Hoffman et al. (2019) show that from 2010 to 2012, only about 1.9

Table 5
Understanding of SSDI work rules.

	One year after enrollment	Two years after enrollment
Panel A. Treatment group members		
Understand that there is no Trial Work Period	34.0	34.9
Understand the termination rules	34.7	33.9
Understand that benefits can be reduced if monthly earnings are sufficiently high	49.0	46.1
Number of responses	2,635	5,054
Panel B. Control group members		
Understand that there is a Trial Work Period	28.0	28.0
Understand that benefits can be terminated if earnings are too high	44.0	43.6
Number of responses	1,438	2,803

Source: POD two-year follow-up survey.
Note: The following three questions assessed the understanding of treatment group members about POD rules: (1) Under POD, do you have a TWP where your benefits remain unchanged regardless of your earnings? (2) Under the POD rules, do your benefits ever terminate if your earnings are too high? (3) Under POD, are your benefits reduced at any time if your monthly earnings are above a level that SSA set for POD? The correct answer to the second question differed depending on the version of the POD rules. The following two questions assessed the understanding of current SSDI rules by control group members: (1) Under current SSDI rules, do you have a Trial Work Period where your benefits remain unchanged regardless of your earnings? (2) Under current SSDI rules, do your benefits ever terminate if your earnings are too high? Beneficiaries had the option to answer that they did not know the answer.

[28] These provisions might be valuable work incentives for beneficiaries as they allow people to test work before benefits can become suspended.

percent of SSDI beneficiaries faced overpayments, but for those who did, the median overpayment amount was more than $9,000 (which was then owed back to SSA). The large amount stems from the fact that monthly SSDI benefit payments are all or nothing – if the benefit is overpaid, it is by definition overpaid by the full benefit amount. By introducing a benefit offset, where beneficiaries can still be paid a partial benefit in a month, the new rules dramatically limited the size of these overpayments.

However, other types of learning costs might have contributed to the limited impacts. Though the new rules reduced the amount of administrative burden, the key question is whether it reduced administrative burden by *enough*, rather than if it reduced administrative burden *at all*. Under the new benefit offset rules, learning costs remained for beneficiaries: POD treatment group members also struggled to understand the new offset rules. Many also struggled with filing monthly earnings on time (a type of compliance cost), which led to overpayments. Below, we provide evidence on these factors.

One important challenge was that treatment group members had substantial difficulty understanding the rules. In the two-year follow-up survey, only 46 percent of treatment group members correctly understood that the benefit offset reduced their monthly benefits for earnings above a key threshold (Table 5, Panel A).[29] Only 34 percent correctly identified whether their benefits could be terminated if their earnings were too high.[30] Understanding of the rules was approximately similar in surveys conducted both one and two years after enrollment, indicating that despite opportunities to further engage with the new rules, knowledge did not improve. Perhaps not surprisingly, offset users were more likely than those who never used the benefit offset to correctly answer those questions. Nonetheless, a particular challenge in this context may relate to the fact that all POD enrollees had some exposure to current rules before enrolling in the demonstration. Thus, knowing how the *change* in the rules might influence the *change* in optimal earnings behavior might be cognitively challenging.

A second important challenge was that people in the treatment group who used the benefit offset faced frequent improper payments, which may have exacerbated issues around understanding how work activity related to adjustments in benefits. Among those who used the offset in 2019, 86 percent had an improper payment.[31] About 74 percent had an overpayment, and about 40 percent had an underpayment. Because about 30 percent of treatment group members used the benefit offset (discussed below), this means more than one in five treatment group members experienced an overpayment. The high prevalence of improper payments may make it challenging for beneficiaries to correctly anticipate the way that their work activity will affect their benefits because it attenuates the connection between earnings and benefit amounts. Recent research by Anand et al. (2022) found that these overpayments significantly reduce work activity. An important contributor to these overpayments was the need to report earnings monthly in a timely fashion: as discussed next, many people faced challenges with reporting timely, though this improved as the demonstration went on. Overpayments were pervasive and not simply one-time events: offset users had an overpayment in about half (46 percent) of offset months. Among all offset users, the median total overpayment amount was $482, and each offset user with overpayments experienced a median of 2.5 months of overpayments. Thus, though the new rules substantially limited the burden associated with overpayments by leading to much smaller overpayment amounts, a significant burden remained through the high frequency of overpayments.

Finally, the short timespan of the demonstration (two years) may have limited the potential for impacts to emerge. Specifically, Congress mandated that the entirety of the demonstration—including planning, recruitment, implementation, and evaluation—be completed within five years. Given the complexities associated with recruitment (which involved contacting over 400,000 SSDI beneficiaries), the implementation period for everyone who enrolled was limited to a minimum of two years.[32] Two years might have not been sufficient for people to fully adjust to the new rules. For example, treatment group members became better over time at reporting their earnings in a timely fashion (a type of compliance cost): over the first few months, only about 30 to 40 percent reported earnings timely, whereas by the end of the demonstration 60 to 70 percent reported on time. Timely earnings reports are important as they can limit the potential for improper payments. Evidence from the similar BOND demonstration also suggests that the short timespan may be important: the limited impacts in earnings above the annualized SGA amount (similar to our definition of substantive earnings) did not emerge until the third year of the demonstration (Gubits et al. 2018).

B. Size of the incentive

An important question is whether the incentive associated with the offset itself encouraged work. While this offset may be predicted to reduce labor supply for those earning above the cash cliff amount, it nonetheless might not have acted as an incentive for a broader range of beneficiaries. For people whose benefits were in suspense at the time of enrollment, POD would be an important increase in benefits, but one that would be predicted to reduce work activity. For such people, the reintroduction of benefits represents a boost to their income – assuming that leisure is a normal good, that would lead to more consumption of leisure (and less work) because of an income effect. Additionally, the offset ramp would represent a work disincentive because of a substitution effect: instead of taking home $1 for each $1 of income earned, such beneficiaries would now only take home $0.50 for each $1 of income earned with the 2-for-1 offset.[33] That in turn decreases the (relative)

[29] The precise survey question asked respondents "Under POD, do you have a TWP where your benefits remain unchanged regardless of your earnings?" Beneficiaries had the option to respond that they did not know the answer, in which case they are considered to not correctly understand the concept.

[30] The precise survey question asked respondents "Under the POD rules, do your benefits ever terminate if your earnings are too high?" The correct answer depended on the version of the POD rules. Beneficiaries had the option to respond that they did not know the answer, in which case they are considered to not correctly understand the concept.

[31] Our analysis of improper payments focuses on 2019, the first full calendar year in which beneficiaries were exposed to the benefit offset, as information on improper payments in 2020 was unavailable at the time of the analysis. Additionally, during the COVID-19 pandemic, SSA changed the way that it administered improper payments, particularly not collecting overpayments, which would make interpretation of findings in that year tricky.

[32] The first mailings went out in January 2018. The first beneficiaries enrolled in February 2018. To be eligible for the demonstration, one had to submit their enrollment paperwork by December 31, 2018. Given the annual nature of earnings data available, this means that we can only focus on 2019 and 2020 as the two full years of implementation.

[33] It is notable to observe the difference in average and marginal tax rates for people with earnings sufficiently high to have their benefits suspended. Consider someone earning $20,000, well above the annual threshold of $15,720 to have benefits suspended for exceeding the SGA amount. For people who are single, with a standard deduction in 2021 of $12,550, their income would lead them to land in the lowest tax bracket of 10% (which covers income up to $9,950). Thus, the marginal tax rate under POD would be 60%, as opposed to 10% under current rules. Yet the average tax rate is far lower under POD than under current rules because of the lost SSDI benefit. Under POD, their average tax rate would be 25.5%: their taxable income of $7,450 would be taxed at a 10% rate, for a total income tax of $745, plus a $4,356 annual reduction in their SSDI benefit (each month, their SSDI benefit would be reduced by $363 as their monthly income of $1,666 exceeds the $940 threshold by $726). In contrast, under current rules, they would pay the same income tax but also lose their full SSDI benefit. Assuming an SSDI benefit amount of $1,035 (the average for those who enrolled in the demonstration), their average tax rate would be 65.8% under current rules.

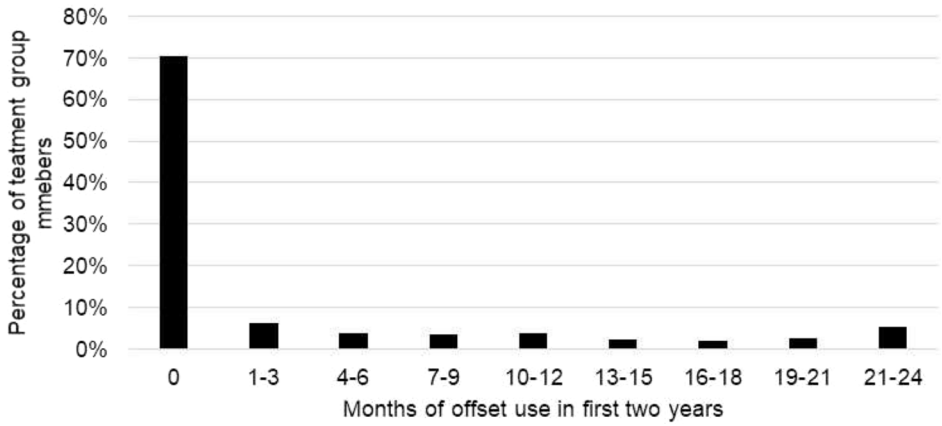

Fig. 2. Distribution of months of offset use in first two years .
Source: Authors' calculations using POD Implementation Data System. Note: Includes all 6,700 treatment group members

price of leisure, thus incentivizing more leisure and less work.

As noted in Section V.A, Table 3 showed that those whose benefits were suspended at the time of enrollment—a group for whom the benefit offset was predicted to reduce work activity—did not change their work activity. In Appendix Tables 1 and 2, we considered alternative definitions to broaden the group for whom the benefit offset might be predicted to reduce work: those having earnings above the SGA amount for at least one month in 2017 (812 people; Appendix Table 2) and those having earnings above the SGA amount in at least one month prior to enrolling since 2014 (1,504 people; Appendix Table 3). These groups capture people for whom benefits increased when the cash cliff was replaced with an offset ramp – all three subgroups have large positive impacts on SSDI benefit amounts that are also differentially significant from the other subgroups that had more limited work activity. Yet there is no statistically significant differential effect on average annual earnings or substantive earnings for any of the three subgroups we consider. The point estimate from our primary estimated impact on average annual earnings (in Table 3) is also positive.

Other groups may also not have faced large incentives for work. As discussed in Section 2.B, people with earnings above the POD threshold ($940 in 2021) but below the SGA amount ($1,310 in 2021 for non-blind beneficiaries) had an ambiguous prediction as to their work activity. A substitution effect would likely lead to reduced work activity: under current rules, additional earnings lead to no change in benefits (since benefits only change when earnings exceed the SGA amount), whereas with the benefit offset, people with earnings in this range would face an additional 50 percent marginal tax rate on earnings given the $1 for $2 offset.[34] But the income effect stemming from a lower benefit payment would likely lead to greater work activity. Consistent with this, we found no change in employment for those who had earnings in this range before enrolling in the demonstration. Additionally, within the first two years after enrolling, approximately 30 percent of treatment group members had earnings sufficiently high to have used the benefit offset (Fig. 2). About 28 percent of these offset users had an offset amount sufficiently low that they experienced reductions in total income relative to current rules.[35] The implied number of people for whom this ambiguous predicted effect applies is thus roughly the same as the number in this subgroup based on earnings prior to enrolling (shown in the second column of Table 3).

Finally, as discussed above and shown in Table 4, there were limited impacts throughout the earnings distribution. We considered earnings above $0, $1,000, about $15,000 (substantive earnings, shown in Table 2), about $30,000, and about $45,000. All of these estimated impacts throughout the earnings distribution are small and not significant (once adjusting for multiple comparisons). Together, these findings offer additional support that the incentive may not have been sufficiently large to influence work activity.

C. Individual and systemic barriers to employment for people with disabilities

A final factor that might have contributed to the limited impacts observed in our study is the individual and systemic barriers to employment faced by people with disabilities. At the systemic level, people with disabilities face discrimination in the labor market. Correspondence studies like Bellemare et al. (2023) and Ameri et al. (2018) have found that people with disabilities are less likely to get a call back for a job interview. Livermore and Honeycutt (2015) highlight the lower rates of employment for people with disabilities as compared to those without disabilities, while Sundar et al. (2018) document some of the broader barriers that people with disabilities face in employment. Further, the large macroeconomic shock associated with COVID-19 during the second year of the demonstration may have differentially hurt people with disabilities (as it did during the Great Recession; Livermore and Honeycutt, 2015).

Our data point to the limited work capacity of individual SSDI beneficiaries specifically. About 90 percent of people in the baseline survey agreed that it was difficult to work because of a physical or mental condition. People also cited other reasons, such as limited skills or a fear of losing benefits or health insurance, as factors that made it difficult to work. Additionally, in the control group, about 55 percent of people did not have any earnings during the two year study period (Table 4). About two-thirds reported not being employed in the past year (Table 4), while only ten percent earned above roughly $15,000 per year (the threshold for substantive earnings in Table 2). Thus, the substantial disabilities that people must have to qualify for SSDI—defined as an inability to

[34] In contrast to the example noted above for those with benefits suspended, both marginal and average tax rates are higher under POD rules than current rules for people in this range. Here, we consider earnings of $13,500, the midpoint between the annualized POD threshold and SGA amounts. Similar to someone whose benefits are suspended, the marginal tax rate under POD would be 60%, as opposed to 10% under current rules based on falling in the lowest tax bracket. Their average tax rate under POD would be 8.9%: their taxable income of $950 would be taxed at a 10% rate, for a total income tax of $95, plus a $1,110 annual reduction in their SSDI benefit (each month, their SSDI benefit would be reduced by $92.5 as their monthly income of $1,125 exceeds the $940 threshold by $185). In contrast, under current rules, their average tax rate would be only 0.7% (from the $95 income tax paid).

[35] The gap between the TWP and SGA amounts was $350 in 2020, so anyone with an offset amount less than $175 would have had sufficiently low earnings to fall in this range where POD adversely affected benefits.

perform any substantial gainful activity—make it difficult for people to work. This limited work capacity is particularly notable given that people who enrolled in the demonstration had a higher predisposition towards work than those who did not enroll. Yet it is important to note that prior research, such as Maestas et al. (2013) and Gelber et al. (2017) suggest a lingering work capacity for many SSDI beneficiaries, albeit with limited capacity to earn above the threshold for substantive earnings. Additionally, despite the COVID-19 pandemic, average earnings in the control group were nearly identical in both 2019 ($4,992) and 2020 ($4,916).

7. Discussion

We examined the effect of replacing a cash cliff in SSDI rules, where beneficiaries had benefits reduced to $0 if their earnings exceeded a key threshold by even $1, with a benefit offset that reduced benefits by $1 for every $2 in earnings. By drastically reducing the marginal tax rate that beneficiaries face on earnings above the cash cliff, this change might have offered greater incentive to work for many. However, for some, the new rules might have reduced the incentive to work.

Overall, we found no impact on earnings, SSDI benefit amounts, or total income. Our estimates are sufficiently precise to rule out substantive changes for the average volunteer: our estimated 95 percent confidence intervals rule out changes of more than 4 percent for benefit amounts and total income, and changes of more than 10 percent for earnings. Our results around earning above the substantial gainful activity threshold (annually about $15,000) are consistent with other papers like Maestas et al. (2013) and French and Song (2014) that show limited capacity for SSDI beneficiaries to earn at this relatively high threshold. Such findings reinforce the difficulties that policymakers face in trying to induce a large earnings response that can help significantly combat the disadvantages SSDI beneficiaries face in the labor market.

A natural question is why this rule change did not impact beneficiary work activity. We present evidence surrounding three potential mechanisms, though caution that this evidence is more speculative than definitive as it relies on descriptive analyses. First, we show that though the demonstration simplified program rules and reduced administrative burden, administrative burden remained wherein people faced challenges with learning costs: limited understanding of the rules, potentially exacerbated by frequent improper payments when beneficiaries work, may have inhibited work activity. People may also have not had enough time to change their work behavior given the demonstration only ran for two years (the second year also coincided with the start of the COVID-19 pandemic). Second, the incentive may not have encouraged work for people who had previously had benefits suspended because of both income and substitution effects that predicted reduced labor supply. Finally, broader barriers to employment at both the individual-level and systemic-level may also have prevented people from working. Each of these three elements likely play a role in why replacing the cash cliff with a benefit offset as tested in POD had no effect on employment-related outcomes.

Our findings around administrative burden are important for informing policy. Much literature around administrative burden has focused on the ways that administrative burdens, such as through complex application processes, deter and limit participation in programs (e.g., Deshpande and Li 2019; Brodkin and Majmundar 2010; Finkelstein and Notowidigdo 2019; Moynihan et al., 2016). Yet the same logic might also dictate that administrative burden can make using the programs more challenging among those qualified (e.g., Homonoff and Somerville 2021; Moynihan, Herd, and Harvey 2015; Heinrich et al., 2022). Our results do not provide a definitive explanation for the POD findings but fit with the framework developed by Barnes (2021) that highlights redemption costs in understanding challenges associated with using benefits. In particular, the analysis of improper payments, as well as the lack of knowledge of program rules, indicate that administrative burden likely affected POD participants' behavioral responses. One way to potentially address these challenges in a future demonstration similar to POD would be to use only new beneficiaries. New beneficiaries would not face the same adjustment costs that beneficiaries in POD faced—things like adjusting to a new expectation regarding earnings reporting to minimize overpayments and adjusting to updated program rules (which required both understanding the existing rules and the new rules). New beneficiaries would have a clean slate, making the various rule simplifications potentially easier for beneficiaries to parse.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix Table 1. . Baseline characteristics of POD enrollees compared with non-volunteers

Variable	POD enrollees	Non-volunteers	*p*-value
Sex and age			
Female	55.0	49.5	0.000
Mean age (years)	46.5	47.3	0.000
Primary diagnosis			
Neoplasms	2.9	3.3	0.000
Mental disorders	38.4	33.1	
Intellectual disabilities	2.6	3.7	
Back or musculoskeletal system	20.2	24.5	
Nervous system disorders	6.3	7.4	
Circulatory system disorders	5.8	6.5	
Genitourinary system disorders	4.2	3.2	
Injuries	3.8	4.1	
Respiratory	1.7	1.7	
Several visual impairments	2.3	2.7	
Digestive system	1.5	1.8	
Other impairments	10.3	8.2	
Beneficiary program characteristics			
Mean SSDI duration (months)	103.9	102.1	0.031
Monthly SSDI benefits ($)	1,035	1,129	0.000
Has representative payee	6.9	12.9	0.000
Concurrent SSI receipt	18.2	14.6	0.000
Employment history			
History of TWP-level earnings since 2014	19.2	7.9	0.000

(*continued on next page*)

(*continued*)

Variable	POD enrollees	Non-volunteers	*p*-value
History of SGA-level earnings since 2014	14.9	5.9	0.000
Had a Ticket assigned in last four years	12.5	4.4	0.000
Observations	10,070	409,344	

Source: Authors' calculations using SSA program records and POD Implementation Data System.

Note: Unless otherwise noted, all table entries are percentages. The *p*-values in the final column compare means for continuous variables, proportions for binary variables, and the distribution for the primary diagnosis.

Appendix Table 2. . Impacts on primary outcome, by 2017 earnings status at time of enrolment

	(1) Average annual earnings	(2) Substantive earnings	(3) SSDI benefit amount	(4) Average annual income
Panel A. Any months in 2017 with SGA-level earnings				
Treatment	−695	−1.1	1,188**+	493
	(1,286)	(3.3)	(484)	(1,160)
Control mean	17,475	40.6	8,022	25,322
Observations	812	812	812	812
Panel B. No months in 2017 with SGA-level earnings				
Treatment	234	1.4***+	−6	241
	(180)	(0.6)	(105)	(185)
Control mean	3,686	6.9	12,115	15,659
Observations	9,258	9,258	9,258	9,258
Panel C. Differential impact *p*-value				
Treatment	0.474	0.438	0.016	0.830

Note: The number for treatment represents an estimate of β from equation (1), representing the estimated impact of being assigned to the treatment group, for those within each subgroup. Standard errors, reported in parentheses, are robust to heteroscedasticity. The differential impact *p*-values in Panel C come from a test of whether the impact estimate for those had any months in 2017 with SGA-level earnings is equal to the impact estimate for those who had no months in 2017 with SGA-level earnings. All monetary values are in 2019 dollars. Substantive earnings is an indicator variable for whether the beneficiaries' average earnings over 2019 and 2020 exceeds the annualized SGA amount over the full two-year period. All outcomes are measured over the 2019 and 2020 calendar years. The exception is that SSDI benefit amounts are measured over the two years after POD enrollment.

***/**/* indicate estimate is significantly different from 0 at the 1/5/10 percent level using unadjusted p-values.

+++/++/+ indicate estimate is significantly different from 0 at the 1/5/10 percent level after adjusting for multiple hypothesis testing using sharpened false discovery rate q-values.

Appendix Table 3. . Impacts on primary outcome, by recent earnings status at time of enrolment

	(1) Average annual earnings	(2) Substantive earnings	(3) SSDI benefit amount	(4) Average annual income
Panel A. Any months since 2014 with SGA-level earnings				
Treatment	−620	0.8	1,160***+++	599
	(866)	(2.3)	(342)	(772)
Control mean	13,985	30.6	9,525	23,239
Observations	1,504	1,504	1,504	1,504
Panel B. No months since 2014 with SGA-level earnings				
Treatment	206	1.1**	−50	158
	(175)	(0.5)	(108)	(185)
Control mean	3,321	6.3	12,139	15,338
Observations	8,566	8,566	8,566	8,566
Panel C. Differential impact *p*-value				
Treatment	0.349	0.912	0.001	0.578

Note: The number for treatment represents an estimate of β from equation (1), representing the estimated impact of being assigned to the treatment group, for those within each subgroup. Standard errors, reported in parentheses, are robust to heteroscedasticity. The differential impact *p*-values in Panel C come from a test of whether the impact estimate for those had any months since 2014 with SGA-level earnings is equal to the impact estimate for those who had no months since 2014 with SGA-level earnings. All monetary values are in 2019 dollars. Substantive earnings is an indicator variable for whether the beneficiaries' average earnings over 2019 and 2020 exceeds the annualized SGA amount over the full two-year period. All outcomes are measured over the 2019 and 2020 calendar years. The exception is that SSDI benefit amounts are measured over the two years after POD enrollment.

***/**/* indicate estimate is significantly different from 0 at the 1/5/10 percent level using unadjusted p-values.

+++/++/+ indicate estimate is significantly different from 0 at the 1/5/10 percent level after adjusting for multiple hypothesis testing using sharpened false discovery rate q-values.

Data availability

The data that has been used is confidential.

References

Ameri, M., Schur, L., Adya, M., Bentley, F.S., McKay, P., Kruse, D., 2018. The disability employment puzzle: A field experiment on employer hiring behavior. ILR Rev. 71 (2), 329–364.

Anand, P., Hoffman, D., Jones, J.T., Lukashanets, S., 2022. Labor supply response to overpayment notifications: evidence from social security disability insurance. *Contemp. Econ. Policy*. 40 (2), 304–322.

Anderson, M.L., 2008. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *J. Am. Stat. Assoc.* 103 (484), 1481–1495.

Barnes, C.Y., 2021. It takes a while to get used to": the costs of redeeming public benefits. *J. Public Adm. Res. Theory* 31 (2), 295–310.

Bellemare, C., Goussé, M., Lacroix, G., Marchand, S., 2023. Physical disability and labor market discrimination: evidence from a video résumé field experiment. *Am. Econ. J. Appl. Econ.* 15 (4), 452–476.

Brodkin, E.Z., Majmundar, M., 2010. Administrative exclusion: organizations and the hidden costs of welfare claiming. *J. Public Adm. Res. Theory* 20 (4), 827–848.

Chen, S., Van der Klaauw, W., 2008. The work disincentive effects of the disability insurance program in the 1990s. J. Econom 142 (2), 757–778.

Deshpande, M., Li, Y., 2019. Who is screened out? Application costs and the targeting of disability programs. *Am. Econ. J. Econ. Pol.* 11 (4), 213–248.

Finkelstein, A., Notowidigdo, M.J., 2019. Take-up and targeting: experimental evidence from SNAP. *Q. J. Econ.* 134 (3), 1505–1556.

French, E., Song, J., 2014. The effect of disability insurance receipt on labor supply. *Am. Econ. J. Econ. Pol.* 6 (2), 291–337.

Gelber, A., Moore, T.J., Strand, A., 2017. The effect of disability insurance payments on beneficiaries' earnings. *Am. Econ. J. Econ. Pol.* 9 (3), 229–261.

Gubits, Daniel, David Stapleton, Stephen Bell, Michelle Wood, Denise Hoffman, Sarah Croake, David R. Mann, et al. 2018. "BOND Implementation and Evaluation: Final Evaluation Report." Final Report submitted to the Social Security Administration. Cambridge, MA: Abt Associates, and Washington, DC: Mathematica Policy Research.

Guess, Andrew M., Neil Malhotra, Jennifer Pan, Pablo Barberá, Hunt Allcott, Taylor Brown, Adriana Crespo-Tenorio et al. "How do social media feed algorithms affect attitudes and behavior in an election campaign?" Science 381, no. 6656 (2023): 398–404.

Hays, Ron D., Cathy D. Sherbourne, and Rebecca Mazel. 1995. "User's Manual for the Medical Outcomes Study (MOS) Core Measures of Health-Related Quality of Life." Santa Monica, CA: RAND Corporation. Available at https://www.rand.org/pubs/monograph_reports/MR162.html.

Heinrich, C.J., Camacho, S., Henderson, S.C., Hernández, M., Joshi, E., 2022. Consequences of administrative burden for social safety nets that support the healthy development of children. *J. Policy Anal. Manage.* 12 (7), 107–132.

Herd, Pamela, Moynihan, Donald P., 2019. Administrative burden: Policymaking by other means. Russell Sage Foundation.

Hock, Heinrich, David Wittenburg, Michael Levere, Noelle Denny-Brown, Heather Gordon, 2020. Promoting opportunity demonstration: recruitment and random assignment report. Final report submitted to the Social Security Administration. Washington, DC: Mathematica.

Hoffman, D., Fischer, B., Jones, J.T., McGuirk, A., Loewenberg, M., 2019. Work-related overpayments to social security disability insurance beneficiaries: prevalence and descriptive statistics. *Soc. Secur. Bull.* 79 (2), 65–83.

Homonoff, T., Somerville, J., 2021. Program recertification costs: Evidence from SNAP. *Am. Econ. J. Econ. Pol.* 13 (4), 271–298.

Jolly, N.A., Wagner, K.L., 2023. Work-limiting disabilities and earnings volatility. *Labour Econ.* 81, 102333.

Kostøl, A.R., Mogstad, M., 2014. How financial incentives induce disability insurance recipients to return to work. Am. Econ. Rev. 104 (2), 624–655.

Livermore, G.A., Honeycutt, T.C., 2015. Employment and economic well-being of people with and without disabilities before and after the great recession. J. Disabil. Policy Stud. 26 (2), 70–79.

Maestas, N., Mullen, K.J., Strand, A., 2013. Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt. *Am. Econ. Rev.* 103 (5), 1797–1829.

Moore, T.J., 2015. The employment effects of terminating disability benefits. *J. Public Econ.* 124, 30–43.

Moynihan, D., Herd, P., Harvey, H., 2015. Administrative burden: Learning, psychological, and compliance costs in citizen-state interactions. *J. Public Adm. Res. Theory* 25 (1), 43–69.

Moynihan, D.P., Herd, P., Ribgy, E., 2016. Policymaking by other means: Do states use administrative barriers to limit access to Medicaid? *Adm. Soc.* 48 (4), 497–524.

Nichols, Austin, Jeffrey Hemmeter, Debra Goetz Engler, eds., 2021. Lessons from SSA Demonstrations for Disability Policy and Future Research. Rockville, MD: Abt Press.

Ruh, P., Staubli, S., 2019. Financial incentives and earnings of disability insurance recipients: evidence from a notch design. Am. Econ. J. Econ. Pol. 11 (2), 269–300.

Schimmel, J., Stapleton, D., Song, J., 2011. How common is 'parking' among social security disability insurance beneficiaries? Evidence from the 1999 change in the earnings level of substantial gainful activity. *Soc. Secur. Bull.* 71 (4), 77–92.

Sundar, V., O'Neill, J., Houtenville, A.J., Phillips, K.G., Keirns, T., Smith, A., Katz, E.E., 2018. Striving to work and overcoming barriers: Employment strategies and successes of people with disabilities. J. Voc. Rehab. 48 (1), 93–109.

von Wachter, T., Song, J., Manchester, J., 2011. Trends in employment and earnings of allowed and rejected applicants to the social security disability insurance program. *Am. Econ. Rev.* 101 (7), 3308–3329.

Weathers II, R.R., Hemmeter, J., 2011. The impact of changing financial work incentives on the earnings of social security disability insurance (SSDI) beneficiaries. *J. Policy Anal. Manage.* 30 (4), 708–728.

Wittenburg, David, Michael Levere, Sarah Croake, Stacy Dale, Noelle Denny-Brown, Denise Hoffman, Rosalind Keith, David R. Mann, Shauna Robinson, Rebecca Coughlin, Monica Farid, Heather Gordon, Rachel Holzwart, 2022. Promoting opportunity demonstration: final evaluation report. Final report submitted to the Social Security Administration. Washington, DC: Mathematica.