

Outreach nudges to clients with incomplete returns

Date implemented: November 2022

Date written: January 2023

Abstract: In 2022, 42% of GetCTC clients who started a return did not finish it. This study explores the impact of outreach nudges to these half-finished clients. Of 79,363 clients with incomplete returns, 60% were sent a text message or email nudge to complete their return, including a prompt to call an assistance hotline for more help; 30% got a message without the hotline; and 10% received no message at all. The outreach nudges were successful in generating additional returns, though only marginally: The nudges increased rates of submitting a return by 0.75 percentage points ($p < .001$) and rates of getting an accepted return by 0.2 percentage points ($p = .019$). There is some evidence that the nudges are most effective for clients whose returns are 2-4 months old. While clients in the hotline subtreatment are more likely to call the hotline than those in the no-hotline subtreatment, hotline provision does not appear to have any effect on submission or acceptance rates. The low-touch nudges—without hotline assistance—are cost-effective, generating one additional submission per 132 nudges and one additional acceptance per 500 nudges. However, they are not enough to meaningfully change the GetCTC completion rate. It remains possible—and consistent with limited qualitative evidence—that clients who drop off part-way through the service do not in fact need to file, and that low completion rates are to be expected in a competitive tax filing context. More research would be needed to understand what, if anything, could be done to dramatically increase the GetCTC completion rate.

Authors: Gabriel Zucker, Maximilian Hell

Other experimental results and research from GetCTC 2022 are available [here](#).

1. Research Questions

This study was designed to measure the impact of offering additional nudges and assistance to GetCTC clients who had started but not finished returns as of early November 2022.

The study looked at GetCTC clients using standard simplified filing, rather than EITC functionality. A similar study looking specifically at the EITC functionality [is available here](#).

Since 2021, it has been clear that a significant portion of clients who start using GetCTC do not actually finish the process and submit returns to the IRS. In 2021, of clients who entered their SSN (or ITIN) in

GetCTC, only 64.2% submitted a return. In 2022, this figure was 41.8%.¹ On one hand, these are relatively large numbers; it is common in online applications to lose the overwhelming majority (say, 97%) of users in a long funnel, rather than only 58%. On the other hand, this still means losing a large number of clients. If they could all be made to finish, it could roughly double the impact of simplified filing.

But there are two related questions regarding this group:

- First, is there a substantive reason that these clients dropped off? That is, is drop-off to a degree intentional and indicative of the fact that these clients know that they do not need to file a return (because, for example, they have already done so)?
- Second, even if there is *not* a substantive reason, are nudges actually successful in moving the needle? What about nudges that offer additional assistance?

In 2021, we observationally examined the impact of sending nudges to these clients stuck in the process; results are available in [Lessons from Simplified Filing, 2021](#) p. 95. The non-experimental results were inconclusive, but they did not appear consistent with a significant impact from sending additional nudges. We speculated that the first question above was meaningfully driving the result—that clients who dropped off knew they did not need to file a return, and they did not respond to nudges because they knew they did not need to. These results, though, were not rigorous.

In 2022, we also sought to explore the first question by interviewing clients who dropped off. This is an inherently difficult project, since clients who stop using a service are systematically difficult to reach. A small-n and possibly unrepresentative study of five clients found that four dropped off for a substantive reason: Two realized that they had already gotten all of their tax benefits and had nothing else to claim, and two more realized that someone else was, in fact, entitled to claim (and already had claimed) their children. Only one dropped off for lack of a piece of needed information (a bank account number). If these patterns are reflective of the drop-off population more broadly, it would suggest that most clients dropping off are doing so for a good reason—and therefore, it is unlikely that nudges would be effective in generating more returns.

This study was designed to more rigorously examine the impact of nudges. While the study adds no further detail about the first question outlined above (why did people drop off?), it can offer precise answers on the second (are nudges helpful?).

2. Study Design and Implementation

The sample for this study consisted of clients who, on November 10, 2022, had started a return with GetCTC but had not yet submitted a return. Specifically, we included clients who:

- Had started a return, including entering their SSN or ITIN and a confirmed² contact method (phone or email)
- Had not submitted a return (i.e., did not have an e-file submission)

¹ The number is appreciably higher in [our paper on GetCTC funnel and completion rate](#), where the analysis is restricted to a period that may be more representative of the typical GetCTC experience.

² Meaning that the client entered the phone or email and a six-digit code sent to the contact method.

- Had clicked ‘Don’t claim EITC’ (and were therefore not included in [the related experiment](#) intended to help clients stuck in the EITC flow)

Note that clients were included in the sample regardless of how long ago they may have started the return. As such, there were clients in the sample who had started a return on May 15 and not visited the site since that date, alongside clients who started a return just a week or two earlier and updated it recently. Given the temporal patterns of return completion, this suggests that some members of the sample (those who started more recently) were probably more likely, *ceteris paribus*, to complete their return.

This yielded a group of 79,363 clients. Clients received one of three treatments:

Hotline Nudge	Hello! We see that you started to claim your cash benefits on GetCTC. To finish your application, please call 877-907-0937 or simply log back into your account and answer the remaining questions by Nov 15: https://www.getctc.org/deadline	47,661 (60%)
Non-hotline Nudge	Hello! We see that you started to claim your cash benefits on GetCTC. To finish your application, please log back into your account and answer the remaining questions by Nov 15: https://www.getctc.org/deadline	23,837 (30%)
Control	Did not receive a message	7,947 (10%)

The phone number in the Hotline Nudge group was the tax assistance hotline operated by SimplifyCT. Hotline operators were trained to help clients through many common issues in filing a return.

Messages were sent on November 10. Messages were sent through The Hub, GetCTC’s client interaction system, in the same vein as any other messages sent to clients—including updates on return status and customer service interactions with clients experiencing issues. Clients receive messages as text messages or emails, according to their indicated preference. Messages sent as text messages would appear from Code for America’s shortcode, in the same thread with all other messages from the product.

3. Results

Before looking at the outcome in terms of tax returns, we look first at the process data: Did treatment assignment impact the assistance clients received from the hotline? Table 1 suggests that yes—clients assigned to the hotline status were far more likely to call the hotline than clients in the other groups. Interestingly, the non-hotline group also appears to be more likely to call the hotline than the control group. The nudge seems to have encouraged clients to look around for resources to complete their return, even when those resources were not directly mentioned in the message.

Table 1: Calls to hotline by treatment status

	Msg, Hotline	Msg, No hotline	Control	P(hotline = no hotline)	P(no hotline = control)

Calls to hotline	0.64% (304)	0.06% (14)	0.01% (1)	<.001	.101
------------------	-------------	------------	-----------	-------	------

Notes: Calls to hotline identified by matching phone numbers of assigned clients to hotline call logs. Because clients could opt to call from a different phone number, the numbers are a floor estimate of the number of calls made. Calls are included if they were made after messages were sent.

Table 2 shows the impact of the treatment on actual submission of returns. Column 1 shows that both treatments significantly increase the rate of return submissions. The increase is large in relative terms, though small in absolute terms—only about 0.75 percentage points. Across the 70,000 clients in the treatment groups, in other words, the nudges solicited about 550 additional submissions. Column 2, though, shows that these additional submissions were disproportionately likely to be rejected. The treatment effect is still significant ($p=.022$) but smaller—about 0.2 percentage points. Across the 70,000 clients in the treatment groups, the nudges solicited about 150 additional accepted returns.

Providing the hotline, meanwhile, has no impact on the submission or acceptance rates, relative to a nudge without the hotline.

Table 2: Return submissions by treatment status

	Any return submitted (1)	Any return accepted (2)
Message, with hotline nudge	1.49% (710)	0.50% (239)
Message, no hotline nudge	1.57% (374)	0.52% (124)
Message, pooling with & no hotline	1.52% (1084)	0.51% (363)
Control	0.76% (60)	0.31% (25)
P(message = control)	$p < 0.0001$	$p = 0.019$
P(hotline = no hotline)	$p = 0.413$	$p = 0.740$

Notes: Clients were matched to submissions by their GetCTC Client ID. p -values are from χ^2 tests.

Given that the marginal returns were somewhat more likely to be rejected, it is worth looking at the distribution of rejection reasons across the cohorts. Table 3 shows the breakdown in rejection reasons. The results are ambiguous and do not clearly suggest that the increase in rejections among treatment submissions was more likely attributable to any particular reason.

Table 3: Most common rejection reasons by treatment condition

Reject Reason	Hotline	No Hotline	Control	P(treat = control)	P(hotline = no hotline)
All clients with a reject	538	288	42		
Filer already filed a return	48.33% (260)	39.58% (114)	40.48% (17)	.537	.016

Prior year AGI incorrect	21.38% (115)	24.65% (71)	16.67% (7)	.374	.283
Dependent already claimed	13.20% (71)	19.79% (57)	23.81% (10)	.151	.013
IP PIN missing or incorrect	6.13% (33)	4.51% (13)	11.90% (5)	.089	.333
Filer already claimed	2.42% (13)	2.08% (6)	0.00% (0)	.320	.761
Name SSN mismatch	2.04% (11)	2.78% (8)	0.00% (0)	.320	.503
Spouse AGI missing / incorrect	1.49% (8)	1.04% (3)	0.00% (0)	.452	.595
1095A - file full return	0.93% (5)	0.69% (2)	0.00% (0)	.549	.726

Notes: p -values from χ^2 tests. IRS reject codes broken into categories based on the nature of the reject. Counts of clients with particular rejects are in parentheses; counts don't sum to the total number of clients with a reject because the least common rejects are omitted.

Given the significant heterogeneity in baseline propensity to complete a return based on the age of the return, one might assume the treatment may have heterogeneous effects on clients with older and newer returns. This effect could go either way; perhaps older clients are more disconnected and more likely to benefit from a nudge, or perhaps they are disconnected enough from the product that the nudge itself is not useful.

Table 4 shows the results of regressions controlling for age of return, with an interaction term of treatment and return age. (Treatment is defined as receiving a message, in either the hotline or non-hotline variants.) The results show that the treatment is, indeed, significantly less effective in eliciting submissions for older returns.

Table 4: Heterogeneous treatment effects by age of return

	(1) Submitted	(2) Accepted
Treatment	.0154** (.0035)	.0018 (.0020)
Treatment*Age	-.00007* (.00003)	.000003 (.000018)
Age	-.00003 (.00003)	-.00003 (.00002)
Constant	.0109** (.0033)	.0060** (.0019)
Observations	72,574	72,574
R-squared	0.002	0.000

Notes: Standard errors in parentheses. * $p < .05$, ** $p < .01$. Treatment defined as hotline or non-hotline group.

Appendix A uses splines to explore this heterogeneity graphically and further suggests there may be some non-linearity in the effect. Both very new and very old returns may have smaller treatment effects, with returns around 2-4 months old the most likely to be impacted by the intervention.

4. Discussion

This study provides definitive evidence that nudges to clients with incomplete returns do have a non-zero impact. While these additional marginal returns generated are more likely to be rejected, they are not overwhelmingly so. As such, it takes around 132 such nudges to generate one additional submission and 500 such nudges to generate one additional acceptance. These figures compare relatively favorably, for example, to the impact of cold text message outreach to high-value populations who have *not* started returns—well-implemented cold outreach campaigns to former GetCTC clients or to SNAP and WIC beneficiaries generally generated one submission per 300-500 outgoing messages. As such, the benefit of sending such nudges is in the same order of magnitude of impact as other high-value outreach, and it is probably well worth the cost. There is also suggestive evidence that this outreach is best done about 2-4 months after starting a return.

(Notably, one cost not included in this cost-benefit analysis is the decrease in ability to contact clients in the future if the additional messages precipitate a large number of opt-outs. Especially for a service that is expected to run for many years, this risk should be priced into any cost-benefit calculation.)

On the other hand, there is no real evidence that offering hotline assistance increases the impact of such outreach: There is no difference between filing rates of clients in the hotline and no-hotline subtrements, despite the fact that a significantly higher rate of clients in the hotline subtrement do indeed call the hotline. The results do not provide any evidence that hotline assistance is effective in increasing the completion rates of unfinished returns.

More broadly, it is important to keep these results in perspective: Even in the treatment group, 98.5% of clients in the sample did not submit a return. Relative to the overall completion rate of about 42%, a nudge would increase the completion rate to around 42.5%—not to 50%, or 75%. The paper does not provide any meaningful information about how to *drastically* change the GetCTC completion rate. Even the comparison to cold outreach could be looked at less favorably: It is striking that a message to a generic SNAP beneficiary who may already have filed taxes has the same order of magnitude of effect as a message to someone who already started filing their return with GetCTC and who should, in principle, be a far more promising lead.

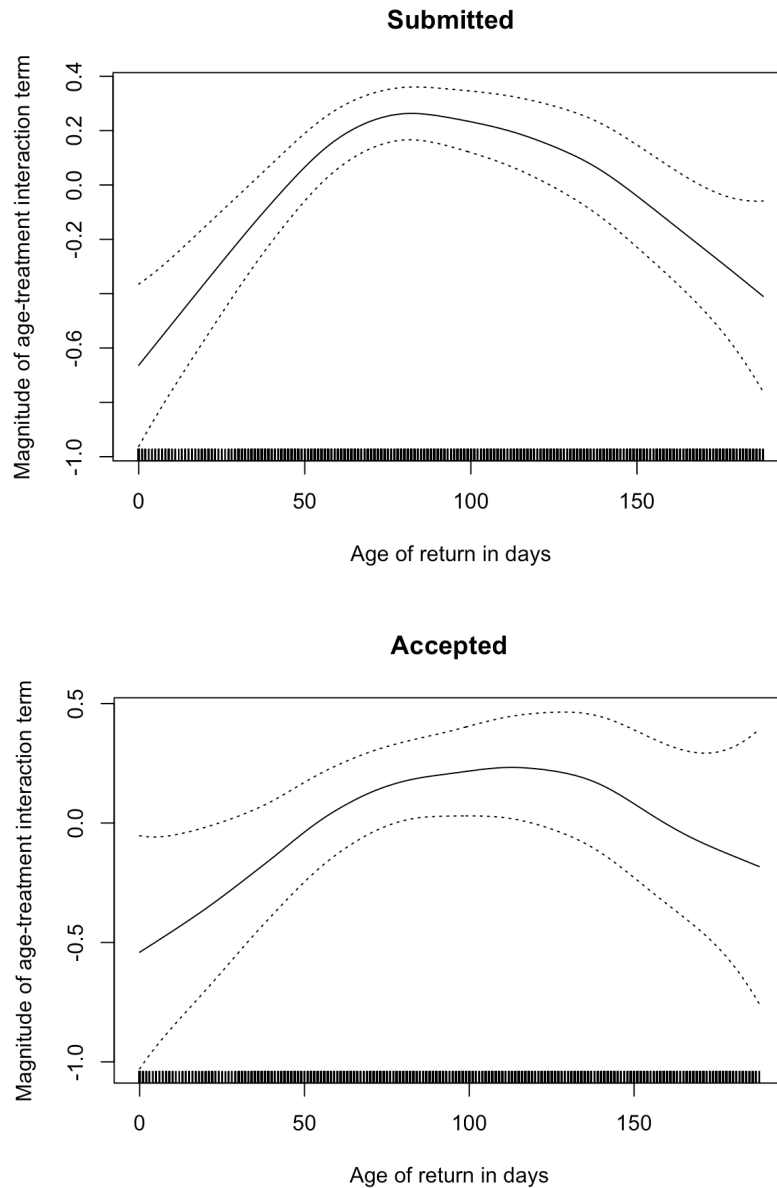
The fact that additional hotline assistance was not effective also suggests—but does not conclusively prove—that lack of assistance may not be the barrier that is preventing clients from completing GetCTC or stopping the nudges from being effective. Of course, it is possible that still more assistance would make the difference, but the results do not provide any evidence that moving in that direction is helpful.

As noted above, this study cannot shed light on the core question of *why* clients are not finishing returns—only whether nudges are effective. One possible explanation for nudges' low effectiveness is that perhaps, as discussed above, half-finished clients do not need to file and know that they do not

need to file. Limited qualitative research from 2022, discussed in Section 1, provides some support for this hypothesis.

Two additional directions of further research could be fruitful in exploring this question. First, it would, in principle, be productive to speak directly with more clients who begin and do not finish returns. Such a research project is fraught, however, since clients who do not complete returns are inherently less connected and less responsive than those who do, making them less likely to respond to a research request. Any work done in this vein can suffer from selection bias. Second, it would be valuable to cross-reference other IRS data regarding half-finished clients. What fraction of the clients who do not complete returns went on to file using a different service instead? What fraction had already filed at the point that they started their return on GetCTC? The answers to these questions could help explain the reasons for drop-off, and they could point either to policy solutions to reduce drop-off or an understanding that the drop-off itself is, paradoxically, intended behavior.

Appendix A. Non-linearity in age heterogeneity



Estimates from a generalized logistic model including age, treatment status, and age-treatment interaction term to predict return submissions and return acceptances. Estimates are fitted using the gam package in R.