Health Economics

WILEY

Health Economics

SHORT RESEARCH ARTICLE OPEN ACCESS

Health Economics Letters

Promoting Safe Opioid Disposal: Experimental Evidence on Behavioral Messaging With Financial Incentives

Mattie Toma[1] | Russell Burnett[2] | Pompa Debroy[3] | Eugen Dimant[4] | Jean Liu[5] | Elana Safran[3] | Uzaib Saya[6] | Bill Schultz[3]

[1]University of Warwick, Coventry, UK | [2]Department of Justice, Washington, DC, USA | [3]Office of Evaluation Sciences, Washington, DC, USA | [4]University of Pennsylvania, Philadelphia, Pennsylvania, USA | [5]Department of Veterans Affairs, Washington, DC, USA | [6]Department of State, Washington, DC, USA

Correspondence: Mattie Toma (Mattie.Toma@warwick.ac.uk)

Received: 12 May 2025 | **Revised:** 12 January 2026 | **Accepted:** 26 January 2026

Keywords: monetary incentives | nudge | opioids | RCT | reminders

ABSTRACT

Some patients prescribed opioid pills fail to properly dispose of their unused pills, posing a risk for others. We report results from a pre-registered field experiment testing whether behaviorally informed reminder cards increase participation in a financial incentive program for returning unused opioid pills among U.S. Veterans. The cards incorporate principles from behavioral economics, including timely reminders, implementation intention prompts, and loss framing. Relative to a cash incentive alone, the reminder cards increase both the likelihood that patients return unused pills (the extensive margin) and the number of pills returned (the intensive margin). The intervention also improves cost-effectiveness, reducing both the cost per pill returned and the cost per participant who returns pills. These findings show that low-cost behavioral messaging can meaningfully enhance the effectiveness of existing financial incentive programs for safe opioid disposal.

JEL Classification: C93, D9, I12, I38

1 | Introduction

The opioid epidemic in the United States has had profound political, social, and economic consequences, including increased mortality, rising healthcare costs, and reduced workforce productivity (Maclean et al. 2022; Alpert et al. 2022; Alpert et al. 2023; Arteaga and Barone 2024). These effects have been compounded by a dramatic increase in overdose deaths: Over the last 20 years, over half a million people have died of overdoses caused by the nationwide opioid crisis (CDC 2020), with overdose deaths increasing over time (CDC 2023).

A growing understanding of the epidemic's roots has led many policymakers, researchers, and practitioners to turn toward identifying prevention and mitigation strategies (Ruhm 2019; Cutler and Glaeser 2021; Humphreys et al. 2022; Cutler and Donahoe 2024). However, despite substantial policy attention and research investment, efforts to address the crisis have produced mixed results (Alpert et al. 2018; Buchmueller and Carey 2018; Eichmeyer and Zhang 2022; Evans et al. 2019; Maclean et al. 2021).

We contribute evidence from a pre-registered field experiment that evaluates a behavioral intervention aiming to limit the scope for opioid misuse. The intervention is designed to augment the effects of a monetary incentive to motivate patients (Veterans) prescribed opioids for post-surgery pain management to return their unused pills. While prior work shows that financial incentives can motivate opioid returns (Liu et al. 2021), less is known about whether low-cost behavioral messaging can further enhance the effectiveness and cost-efficiency of such programs. Our study directly addresses this gap.

The vast majority of individuals prescribed opioids keep unused pills (Dollar et al. 2022; Gregorian et al. 2020).[1] Unused prescription opioid pills that remain in medicine cabinets, on countertops, and in drawers when they are no longer needed for their prescribed purpose can impact the individual, household members, and their communities by leading to both prescription and illicit opioid misuse and abuse. According to 2023 data (Substance Abuse and Mental Health Services Administration 2024), 39% of individuals 12 years or older who are current misusers of prescription pain relievers self-report having obtained the pain medication from friends or family with or without their knowledge, pointing to the importance of proper disposal for society more broadly. Still, the risks of improper disposal are especially relevant for the patient herself: 47% of misusers report getting the pills directly through prescriptions from healthcare providers. As one preventative measure that aims to decrease the supply of unused prescription opioids in homes and communities, the FDA encourages patients to make use of drug take-back locations to safely dispose of unused or expired medicine (FDA 2023).

The intervention we study was introduced at the White River Junction Veterans Affairs (VA) Medical Center in Vermont, which had recently launched a drug take-back program, "Cash for Your Stash" (Liu et al. 2020, 2021). In this program, patients who returned unused opioid pills to the VA pharmacy (within a 60-day eligibility window following their procedure) received $5 per pill, up to a maximum of $50. Our study tests a behaviorally-informed message aimed at augmenting the effects of the Cash for Your Stash program, building on prior research pointing to the value of interacting low-cost reminders with financial incentives (Karlan et al. 2016; Banerjee et al. 2021).

Specifically, we randomly assigned 512 patients at the VA Medical Center to either a control group eligible for the existing cash buyback program or to a treatment group that additionally received two cards with messages encouraging participation in the program. Messages on the cards incorporated several insights from behavioral economics, including timely delivery of information, an implementation intention prompt, and a loss framing designed to activate the endowment effect.[2] One card was given to patients when they received their opioid pills. Another card was sent to the patient about a week later, when they were likely to be out of the post-op timeframe that requires opioids and therefore know if they had unused pills to return for proper disposal. The efficacy of the intervention is measured using administrative data from VA electronic health records.

Overall, our treatment notably improves outcomes. We find that patients assigned to receive these cards are more likely to return their unused opioids, both at the extensive margin (likelihood of returning pills) and at the intensive margin (number of pills returned). Crucially, we also find that our intervention improves cost-effectiveness: it reduces the cost per pill returned by almost 25% and the cost per person who returns pills by almost 5%.

Our study contributes to an existing literature in behavioral economics aimed at identifying strategies to curb the opioid epidemic. The most successful interventions to date have focused on the *supply side* of the epidemic by targeting the prescription behavior of physicians. For example, randomized controlled trials have tested the impact of physicians receiving letters highlighting the prescribing behaviors of other physicians or fatalities among patients prescribed opioids (see, e.g., Sacarny et al. 2016; Doctor et al. 2018).

We report the results of a randomized controlled trial investigating a complementary approach, targeting the *demand* side of the opioid epidemic. Much of the existing demand-side literature focuses on patients already struggling with opioid misuse (Prendergast et al. 2006; Maclean et al. 2021; Lowenstein et al. 2023). Most directly, our experiment contributes to a nascent literature evaluating interventions to improve the disposal of unused opioids: a recent review of the literature identifies four randomized experiments (two with sample sizes larger than 100) that test the efficacy of interventions such as informing patients of disposal options (Dollar et al. 2022). Another experiment identifies portable tools to improve disposal, testing the effect of providing an activated charcoal bag that allows for at-home disposal (Brummett et al. 2019). Our experiment studies a different type of intervention—behaviorally-informed reminder messages—and is the first to use administrative data rather than survey self-reports to measure disposal outcomes.

More broadly, our experiment allows us to evaluate how behaviorally-informed interventions can enhance social welfare within the operational constraints and diverse patient behaviors encountered in healthcare settings (List 2024). The efficacy of our reminder messages dovetails with other evidence for the benefits of reducing complexity and increasing the salience of actionable decision points in program uptake decisions (Bettinger et al. 2012; Bhargava and Manoli 2015). We are not suggesting that "nudges" in themselves are sufficient to address health crises like the opioid epidemic. However, because the behaviorally-informed messaging utilized in this study meaningfully shifts a high-stakes preventative outcome at a relatively low cost, we believe it represents a "mutually reinforcing" intervention (Chater and Loewenstein 2023) which can be employed alongside more intensive systemic reforms. Reasonable skepticism about the value of behavioral insights is mounting (Maier et al. 2022; Chater and Loewenstein 2023), but there are still settings where leveraging them to update existing policies or incentive structures yields meaningful benefits. Our experiment points to the value of non-monetary behavioral interventions on top of overt monetary incentives.

2 | Study Data and Methods

2.1 | Cash for Your Stash

To encourage patients to return unused opioid pills to the White River Junction VA Medical Center (VAMC), in 2017 the VAMC developed a cash buyback program called "Cash for Your Stash." Ambulatory surgery patients who receive one-time, short-term opioid prescriptions for pain management after outpatient surgery were eligible for the buyback program. In this program, Veterans received $5 per pill, up to a maximum of $50, when they returned unused opioid pills to the VA pharmacy, with a formal 60-day eligibility window. Prior work found that this financial incentive motivated pill return for proper disposal; over a 2-year study period, 30% of patients utilized the program and returned their unused opioid pills (Liu et al. 2021). When providers received data on the pill return rate, opioid prescriptions decreased substantially—62.4% of patients were prescribed an opioid at the beginning of the study period, and this dropped to 38.3% by the end of the study period. Because no corresponding increase in refill rates (a proxy for underprescribing) was observed, we interpret this as suggestive evidence consistent with the broader literature indicating that baseline prescribing is higher than needed (e.g., Eichmeyer and Zhang (2022), Hill et al. (2017), and Barnett et al. (2017) document that the portion of prescribed pills taken and the number of refills is low and that high prescribing rates are associated with long-term opioid dependence). While this analysis is correlational, the fact that information about opioid use may have the potential to shift prescribing behavior further points to miscalibrated beliefs as one potential factor underlying high prescribing rates. Our study builds on the foundations of this existing program.

2.2 | Population

The sample for our project consisted of 512 Veterans visiting the VAMC for outpatient surgeries from June 2019 to January 2022. Consistent with the Veteran population in general, our sample largely consisted of men (91.5%). Patients in our study were on average 61 years old. They came into the clinic for surgeries that fell into several service groups, including: General, Orthopedics, Plastics, and Urology. Following their surgeries, the median patient in our sample was prescribed 10 opioid pills (the 75th percentile was 15 pills prescribed, while some outliers were prescribed as many as 50). The average prescription rate was 11.75 pills. Based on data from our control or "business as usual" condition (reported below), patients returned 2.08 pills on average, in the absence of treatment (or 9.59, on average, among people returning at least one).[3] Note that prescriptions in our sample are overwhelmingly for a particular opioid: 87% are prescriptions for OxyContin, while 13% are for other drugs like Vicodin or Dilaudid. Prescriptions are dispersed across the common procedures in our sample (see Supporting Information S1: Table A.2). Importantly, all prescriptions are for moderate to strong opioids that present significant risks if misused.

With regards to the generalizability of our findings to other patient contexts, the literature suggests that Veterans prescribed in non-VA hospitals and younger adults are prescribed larger opioid quantities on average, with more mixed results for gender (Buys et al. 2025; Gifford et al. 2020; Soffin et al. 2021). This suggests that our treatment effects might represent a lower bound when applied to other contexts with greater scope for pill returns.

2.3 | Experimental Design

We randomly assigned patients eligible for the Cash for Your Stash program to either a control group or a treatment group. Control patients received information about their eligibility for the cash incentive as usual. This information was shared with the patient when their opioid pills were prescribed, through both a sticker on their pill bottle and a printed page detailing the cash buyback option (see Supporting Information S1: Figure A.1). Treatment patients received this standard information plus two reminder cards designed to increase the return of unused pills.[4] We gave the first card (Figure 1) to the patient on the day of the surgery, and mailed the second card to the patient approximately 1 week later (Supporting Information S1: Figure A.2).[5]

The cards included simple and clear action steps, explained the importance of properly disposing of unused pills, and incorporated several features informed by behavioral economics. To activate the endowment effect, we framed the cash incentive as the patient's own money waiting to be claimed (Kahneman et al. 1991; Milkman et al. 2022). We also leveraged implementation intention prompts, encouraging patients to write down a date and make a plan for returning their unused opioid pills to the pharmacy (Milkman et al. 2011). The second card also provided information around the time when patients were able to take action. The cards served to remind patients about the opportunity to return their unused opioid pills. Reminders have been shown to address both present bias and limited memory (Ericson 2017) and have been used in many contexts to motivate health-related behavior change such as increasing gym attendance, educational attainment, flu shots, and COVID-19 vaccinations, among others (see, e.g., Milkman et al. 2011; Altmann and Traxler 2014; Dai et al. 2021; Milkman, Gromet, et al. 2021; Milkman, Patel, et al. 2021; Buttenheim et al. 2022; Milkman et al. 2022; Ruggeri et al. 2023; Duckworth et al. 2025).

Between June 2019 and January 2022, we randomly assigned 258 patients to the treatment group and 254 patients to the control group. Because randomization at the individual level was too burdensome for staff implementing the design, we implemented a clusters-within-blocks randomization procedure: in consecutive 2-week periods (blocks), patients with surgeries in one of the 2 weeks were randomly assigned to the treatment group, while patients with surgeries in the other week were randomly assigned to the control group. On average, there were 4.61 surgeries per week, with a standard deviation of 2.2 (the median is 4, and the inter-quartile range is 3–6).[6]

2.4 | Data

Our analysis relies on existing administrative data from VA electronic health records, which includes the number and timing of any pills that a patient returned. Our primary analyses compare the percentage of patients who returned unused opioid pills between the treatment and control groups, as well as the number returned.

These administrative data allow us to measure a variety of potentially important covariates: the surgical procedure performed; the overall surgery group; the type and amount of opioids prescribed; a patient's age and gender; and whether the prescriber was a resident physician or a physician's assistant rather than an attending physician (which we use as the reference category). We provide summary statistics for all covariates used in our analyses in Supporting Information S1: Appendix C (Table A.1) and summarize the relationship between covariates and pill returns in Supporting Information S1: Appendix E.1 (Figure A.4).

We provide week-level balance estimates for select covariates in Table 1.[7] The one statistically significant difference when looking at covariate-by-covariate comparisons (rates of orthopedic surgeries) is no longer significant after multiple testing adjustment. An additional joint significance test across these 8 covariates using individual-level data does not provide sufficient evidence to reject the null of no average differences with 95% confidence. We adjust for these covariates in our analyses.

FIGURE 1 | Treatment Card 1. We give this to the patient on the day of surgery, along with the standard informational materials. The upper half of this figure shows the front of the card, and the lower half shows the back.

TABLE 1 | Balance on select covariates: cluster-level ($n = 111$).

	T mean	C mean	T count	C count	Adj. Difference	*p*-value
Patient						
Age	61.16	60.80			−0.20	0.88
Female (0/1)	0.10	0.08	21	22	−0.01	0.76
Prescription						
Morphine mg equiv.	89.03	85.13			6.65	0.43
Opioid acute (0/1)	0.02	0.03	4	6	0.00	0.98
Service group						
General (0/1)	0.49	0.49	125	132	−0.05	0.40
Orthopedics (0/1)	0.21	0.15	57	34	0.10	0.03
Plastics (0/1)	0.14	0.16	39	40	0.00	0.96
Urology (0/1)	0.03	0.06	6	13	−0.02	0.28
Joint significance						
F Statistic: 1.371						
p-value: 0.219						

Note: We average select covariates across patients within each of the 111 weeks in our study. For categorical covariates included here, we additionally provide the total count across treatment and control weeks. We then estimate balance separately for each covariate through an OLS regression with block fixed effects and cluster-size weights (HC2 errors). We provide two-sided *p*-values from these individual balance tests for each covariate. None of these imbalances remain significant after multiple testing correction (Holm-Bonferroni adjustment with a 5% family-wise error rate). We also report the results of an individual-level (not week-level) omnibus test of the joint significance of these 8 covariates as predictors of treatment status. Individual-level equivalents of the covariate-by-covariate statistics are similar and support the same conclusions viz. covariate balance.

Across this sample, 27.35% of patients returned at least one pill across our study time frame. Among those returning pills, the median return count was 7 and the average was 9.6 (see Supporting Information S1: Figure A.3 and Table A.3). Additionally, recall that there was a formal 60-day eligibility window for the buyback program. The median patient returning pills did so after 19 days.

2.5 | Empirical Specification

We use OLS linear regressions to compare the rate at which opioid pills were returned and the number of opioid pills returned by patients across our treatment and control groups.[8] Our pre-registered primary outcome is a binary indicator for whether pills were returned. We emphasize findings for the number of pills returned in the main text as well, as changes on the extensive and intensive margins are both relevant for policy. We adjust our treatment effect estimates based on various characteristics of the patients, including their age and gender, their surgical procedure, and the quantity of opioids prescribed (see Supporting Information S1: Appendix C, Table A.1). We also adjust our estimates for the blocked nature of randomization (Gerber and Green 2012).[9] Following our pre-analysis plan, we determine statistical significance using a cut-off point of $p = 0.05$ in a one-sided test alongside a clear directional hypothesis. In this regard, we are following related work (Sacarny et al. 2022) as well as best practices (Lakens 2022) when facing logistical sample size constraints, although our results are robust to using two-sided tests instead.

For our primary specifications, we estimate the following OLS equation:

$$y_i = \alpha + \beta_1 D_i + X_{i,k}\gamma_k + \epsilon_i$$

where i indexes individual patients, y_i stands in for different outcome measures we explore, D_i is a binary indicator for patients whose surgeries were scheduled in a treatment week, and $X_{i,k}$ is a matrix of k covariates defined across patients. These covariates are binary indicators for: different surgery procedures; different surgery groups; the prescriber's identity; different age ranges; prescription strength (morphine milligram equivalent, MME) quartiles; patients who are opioid acute; patients who are women; and the 2-week blocks in which treatment weeks were assigned (block fixed effects). See Supporting Information S1: Appendices C and D for more information on our observed covariates.[10] In all cases where we are controlling for mutually exclusive categories, we leave out one category to be captured by the intercept. Other terms, α, β_1, and γ_k are estimated parameters (in the case of γ_k, a vector of estimated parameters). Our parameter of interest is β_1, the treatment effect estimate.

3 | Study Results

We present our primary results in Table 2 and Figure 2. Panel A of Figure 2 shows that the cards increase the probability that patients return unused opioid pills by 6.88% points over an unadjusted control group mean of 24.8% ($p = 0.007$, one-sided; $p = 0.014$, two-sided). Panel B shows that the cards increase the unused opioid pills returned per person by 1.09 pills, or by 52%, over an unadjusted control mean of 2.08 pills ($p = 0.001$, one-sided; $p = 0.002$, two-sided). Given a median prescription of 10 pills, this suggests that, on average, the cards increase the number of pills returned and properly disposed of from roughly 20% to roughly 30% of those prescribed. Together, these results indicate that the intervention meaningfully increased both the likelihood of returning any pills and the total quantity returned.

To be clear, Panel B's findings in columns 4–6 could be shaped by changes on both the extensive margin (more patients returning at least a few pills) and the intensive margin ("always returners" returning more pills). We therefore consider the effect on the intensive margin effect in isolation in Supporting Information S1: Appendix Table A.13 and find suggestive evidence that the treatment motivated people not just to return pills more often but also to increase the number of pills that they returned.

TABLE 2 | Primary treatment effect estimates.

	(1) Extensive likelihood	(2) Extensive likelihood	(3) Extensive likelihood	(4) Intensive amount	(5) Intensive amount	(6) Intensive amount
Treatment	0.069	**0.069**	0.049	1.093	**1.093**	1.139
	(0.040)	**(0.028)**	(0.030)	(0.493)	**(0.327)**	(0.385)
N	512	512	512	512	512	512
Control mean	0.248	0.248	0.248	2.081	2.081	2.081
Covariates	Yes	Yes	No	Yes	Yes	No
Block FEs	Yes	Yes	Yes	Yes	Yes	Yes
Clustered errors	No	Yes	Yes	No	Yes	Yes
Pre-registered	Yes	No	No	Yes	No	No

Note: Coefficient estimates from OLS regressions, with our preferred specifications in bold (robust standard errors in parentheses). Estimates in models 1 to 3 are probabilities, while the others are counts. Models 1 and 3 (pre-registered) include covariate adjustment and block fixed effects, with HC2 robust standard errors. However, as we assigned treatment at the week level, week-clustered errors may be more appropriate (Abadie et al. 2023). Using a fixed effects regression, we estimate a week-level intra-cluster correlation for the number of pills returned of 21.24%. The remaining models therefore also reports estimates from models that include week-clustered errors. Models 3 and 6 do not include covariate adjustment, which reduces the precision of our estimates. We believe models 2 and 5 provide the most accurate and efficient estimates of our treatment effects, but we provide other specifications in the interests of transparency.

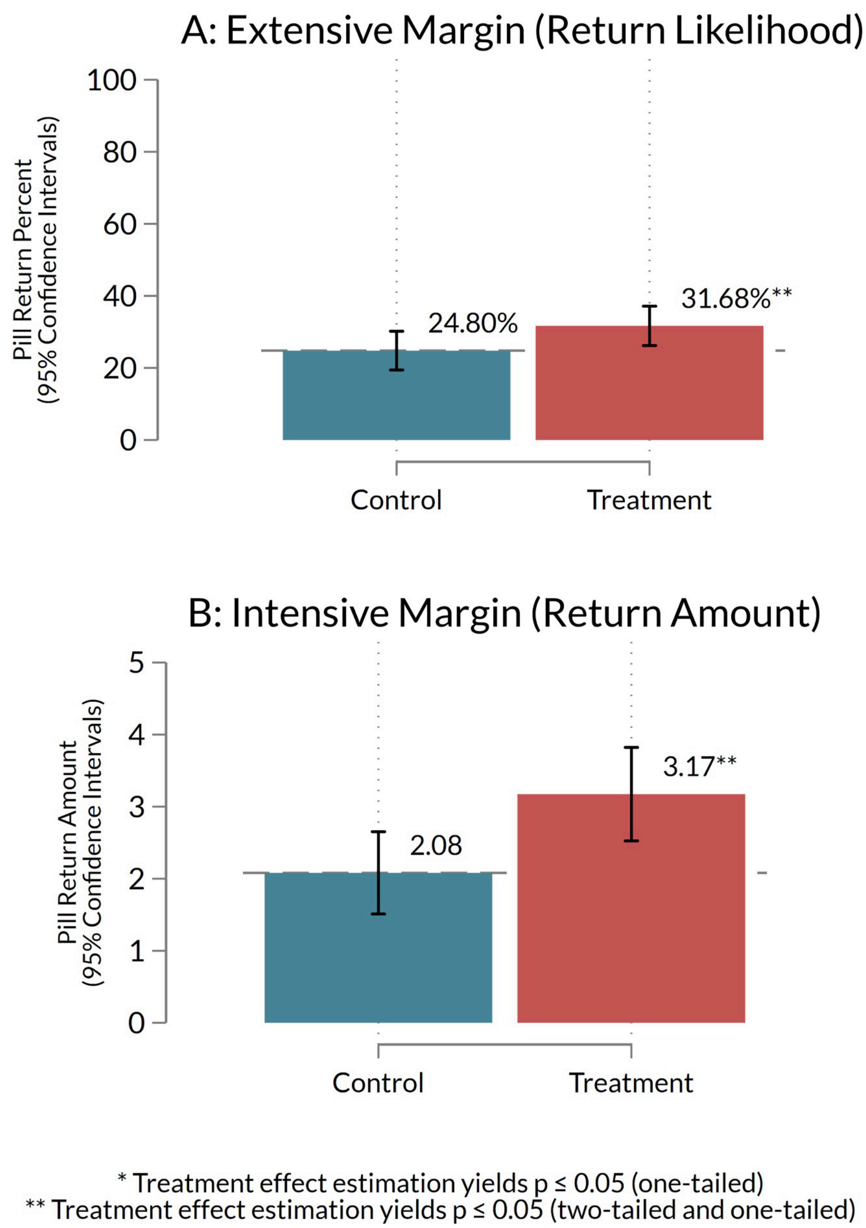

FIGURE 2 | Impacts on pill return likelihood and amount. We estimate the impacts of monetary incentives + behaviorally informed cards (Treatment) compared to the impacts of monetary incentives alone (Control) on pill return likelihood (Panel A) and amount (Panel B). The control point estimates and CIs are based on observed outcome data in the control group. The treatment point estimates add our treatment effects (from Models 2 and five in Table 2) to the observed control group point estimates. The treatment CIs are based on standard errors from those models.

In contrast to these extensive and intensive margin findings, we do not observe a statistically significant effect of the cards on the *time to pill return* (Supporting Information S1: Table A.10). A greater overall return rate in the treatment group could be explained by (1) the cards encouraging more patients to choose to return pills at all, or (2) the cards helping more patients remember to return pills on time. Our results, in conjunction with evidence in Supporting Information S1: Table A.10, therefore provide more consistent support for (1) than (2).

3.1 | Robustness Checks

As noted, the number of patients scheduled for surgery in any 1 week was typically small (an inter-quartile range of 3–6, with a median of 4). With only 512 patients in our sample across 111 weeks, there are reasonable concerns that standard cluster-robust standard error estimators may exhibit downward bias (leading to anti-conservative inferences).

We have aimed to address that concern a few ways. First, we explore the robustness of our findings to the use of alternative cluster-robust variance estimators with better small-sample properties (Pustejovsky and Tipton 2018; MacKinnon et al. 2023). Results in Supporting Information S1: Table A.6, discussed in Supporting Information S1: Appendix E.3, support similar substantive conclusions. Second, we consider an alternative design-based approach to (finite-sample) statistical inference that may also have better small sample properties, both for our balance tests and treatment effect estimates. See

Supporting Information S1: Tables A.8 and A.9 in Appendix E.5. Conclusions from this alternative inference procedure are similar to those that follow from Supporting Information S1: Appendix E.3.

An additional concern is that some limited treatment non-compliance may be leading us to *underestimate* the true impacts of distributing the behaviorally-informed reminder cards. See Supporting Information S1: Appendix E.4 and Table A.7 for estimates from two-stage least squares models that help address this issue. We also note that, as discussed above, there is variation in the opioids patients were prescribed, and it is conceivable that our results could reflect some opioids being returned more often while others were not. Supporting Information S1: Table A.11 shows that the treatment effect among OxyContin prescriptions (most of our sample) is statistically significant and that patients prescribed other opioids (e.g., Vicodin or Dilaudid) see a suggestively larger treatment effect when focusing on the number of pills returned.[11]

Lastly, the Covid-19 pandemic co-occurred with, roughly, the second half of our study time frame. This led to a temporary closure of the clinic (and thus a pause on enrollment into our sample), along with a heterogeneous rescheduling process across surgery types. It also led to some changes in prescription and pill return procedures. Despite this, Supporting Information S1: Table A.12 in Appendix E.8 supports that the effects of our intervention persisted both before and after this period of pandemic-induced changes.

3.2 | Costing

Lastly, we also compared the costs of the baseline "Cash for Your Stash" program with the cost of including the cards.[12] Importantly from a policy perspective, the cards simultaneously increase the number and reduce the cost per pill returned: the total program cost per pill returned is $12.57 in the treatment group, and $15.66 in the control group. The cards also reduce the cost per person returning pills: the total program cost per person returning pills is $124.60 in the treatment group, and $131.20 in the control group. A related question is whether $124.60 is in fact a good price to pay. In 2017, the "monetized burden" for opioid overdose, abuse, and dependence was estimated to be $1.02 trillion (Florence et al. 2021).[13] Researchers accounted for healthcare costs, lost productivity, substance abuse treatment, criminal justice-related costs, reduced quality of life for opioid misuse, and the value of lives lost to fatal overdose. In 2023, 8.9 million people were estimated to suffer from opioid misuse (Substance Abuse and Mental Health Services Administration 2024). Using CPI-U annual averages to adjust the monetized burden estimate for inflation, we estimate an annual cost of opioid misuse of $1.27 trillion, or $143,000 per person in 2023 dollars. As such, given the estimated cost per person returning pills of $124.60, even if only one in one thousand cases of improper disposal led to a case of opioid misuse, the benefits of the program to society would outweigh the costs.

4 | Discussion

We report results from a randomized controlled trial investigating the value-added of behaviorally-informed messaging for encouraging post-surgery opioid return, as a complement to financial incentives. We distributed reminder cards informed by the behavioral economics literature to Veterans at an outpatient surgery facility in Vermont. We demonstrate that these cards increase the return of unused opioid pills for proper disposal on both the extensive and intensive margins (likelihood of returning pills, and the number of pills returned, respectively). We also emphasize the cost savings our intervention produced. Given existing estimates of the social costs of opioid misuse, our figures imply that even a very small reduction in misuse attributable to improved disposal would be sufficient for the intervention's benefits to exceed its costs.

Financial incentives have long been found to serve as an effective means to address the opioid epidemic from the demand side. For instance, contingency management techniques, which involve providing monetary rewards to patients as an incentive to avoid drug use, are found to be among the more effective treatments in promoting abstinence (Prendergast et al. 2006). Additionally, prior work finds that financial incentives motivate Veterans to return unused opioid pills (Liu et al. 2021). We go further, showing that behavioral interventions—in particular, reminder cards—can provide an additional layer of motivation to complement the power of financial incentives. Future work could isolate the relative roles of each of the behavioral mechanisms at play in our bundled treatment. Note that our study cannot speak to concerns over possible "backfiring" effects, where patients given the opportunity to return pills for a cash incentive might respond by foregoing the pain relief they need. We argue that this is less of an issue for the evaluation of improvements to existing buyback programs, and encourage further research on this in the future.

More broadly, our findings have implications for the use of behavioral interventions in policy design. First, our findings recall those of other studies showing that simplifying government communications can improve program take-up outcomes (Bhargava and Manoli 2015). Second, we join a growing number of studies exploring when combining interventions can augment, rather than crowd out, behavior change (Brandon et al. 2019; Fryer et al. 2022; Howley and Ocean 2022). Further work is needed to refine our understanding of the contexts and intervention types that serve as effective complements. Some important related research includes studies on the use of indirect or "meta" nudging to accompany other interventions (Dimant and Shalvi 2022), and studies seeking to learn from evidence of treatment effect heterogeneity (Weiss et al. 2014). Insights from this research may help bolster the effectiveness of behavioral interventions in the future.

Acknowledgments

We thank the entire Office of Evaluation Sciences team as well as numerous seminar participants for helpful feedback. This study's pre-

analysis plan can be found at: https://oes.gsa.gov/assets/analysis/1804-analysis-plan.pdf. The views expressed are those of the authors and do not necessarily reflect the official policy of the US Government, or any particular US Government agency. IRB approval: protocol 1019235.

Funding

This work was supported by the Veterans Health Administration Office of Rural Health and the VHA Innovators Network.

Conflicts of Interest

Included for all authors. Jean Liu reports grants from the Veterans Health Administration Office of Rural Health for patient stipends for the study. No other conflicts of interest are reported.

Data Availability Statement

The data that support the findings of this study are not available due to patient privacy restrictions.

Endnotes

[1] A recent review finds that between one and three-quarters of patients prescribed opioid pills stored them for future use or misplaced them, while only 12%–41% dispose of pills using a method recommended by the U.S. federal government (GAO 2019).

[2] While testing one behavioral feature in isolation would shed light on underlying mechanisms, we opted for the (arguably) maximally-impactful bundled treatment to facilitate the policy-relevant test of whether behavioral insights are able to shift the return of unused opioid pills in this context.

[3] Prescription count is included in the covariate table in Supporting Information S1: Appendix C.

[4] This study has IRB approval from the Veterans Institutional Board of Northern New England (VINNE).

[5] In our dataset, there were no patients who were randomized to receive cards who did not have a mailing address. Our data track whether reminders are sent (see Supporting Information S1: Appendix E.4 for details) but not whether any are returned.

[6] Note that enrollment into our sample was paused between June and August 2021 while we addressed treatment administration issues. All patients during this period had a treatment probability of 0, and so they were excluded from our analysis. See Supporting Information S1: Appendix E.8.

[7] Supporting Information S1: Appendix D shows that a wider variety of patient characteristics are balanced across treatment conditions after adjusting for multiple testing (Supporting Information S1: Tables A.4 and A.5). Supporting Information S1: Appendix B provides additional information about the patient selection procedure.

[8] Our pre-analysis plan can be found on the Office of Evaluation Sciences website.

[9] In practice, we adjust for block fixed effects using a *least squares dummy variable* (LSDV) approach. We note that, since our blocks are 2-week periods, this helps account for spurious temporal trends across our study timeframe. The co-occurrence of the Covid-19 pandemic (and associated policy responses) with the latter half of our study raises special concerns about confounding time shocks. We discuss robustness checks performed with this in mind in Appendix E.8. We are not aware of any other relevant shocks that might coincide with randomly assigned treatment weeks.

[10] All of the models reported in Table 2 include block fixed effects. A few of the models drop the other covariates from $X_{i,k}$.

[11] We consider treatment effect heterogeneity with respect to other covariates as well (Supporting Information S1: Figure A.5).

[12] See Supporting Information S1: Appendix F for a more detailed breakdown of these estimates.

[13] This estimate is for all cases of opioid use disorder, including heroin use, whereas our study is specifically relevant to the misuse of prescription pain relievers. 3.6% of opioid misusers misuse heroin but do not misuse prescription pain relievers (Substance Abuse and Mental Health Services Administration 2024).

References

Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge. 2023. "When Should You Adjust Standard Errors for Clustering?" *Quarterly Journal of Economics* 138, no. 1: 1–35. https://doi.org/10.1093/qje/qjac038.

Alpert, A., W. N. Evans, E. M. J. Lieber, and D. Powell. 2022. "Origins of the Opioid Crisis and Its Enduring Impacts." *Quarterly Journal of Economics* 137, no. 2: 1139–1179. https://doi.org/10.1093/qje/qjab043.

Alpert, A., D. Powell, and R. L. Pacula. 2018. "Supply-Side Drug Policy in the Presence of Substitutes: Evidence From the Introduction of Abuse-Deterrent Opioids." *American Economic Journal: Economic Policy* 10, no. 4: 1–35. https://doi.org/10.1257/pol.20170082.

Alpert, A., S. Schwab, and B. Ukert. 2023. "Opioid Use and Employment Outcomes: Evidence From the U.S. Military." Working Paper.

Altmann, S., and C. Traxler. 2014. "Nudges at the Dentist." *European Economic Review* 72: 19–38. https://doi.org/10.1016/j.euroecorev.2014.07.007.

Arteaga, C. and V. Barone. 2024. "Republican Support and Economic Hardship: The Enduring Effects of the Opioid Epidemic." Working Paper.

Banerjee, A., A. G. Chandrasekhar, S. Dalpath, et al. 2021. *Selecting the Most Effective Nudge: Evidence from a large-scale Experiment on Immunization*. NBER Working Paper.

Barnett, M. L., A. R. Olenski, and A. B. Jena. 2017. "Opioid-Prescribing Patterns of Emergency Physicians and Risk of long-term Use." *New England Journal of Medicine* 376, no. 7: 663–673. https://doi.org/10.1056/nejmsa1610524.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu. 2012. "The Role of Application Assistance and Information in College Decisions: Results From the H&R Block Fafsa Experiment." *Quarterly Journal of Economics* 127, no. 3: 1205–1242. https://doi.org/10.1093/qje/qjs017.

Bhargava, S., and D. Manoli. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence From an Irs Field Experiment." *American Economic Review* 105, no. 11: 3489–3529. https://doi.org/10.1257/aer.20121493.

Brandon, A., J. A. List, R. D. Metcalfe, M. K. Price, and F. Rundhammer. 2019. "Testing for Crowd out in Social Nudges: Evidence From a Natural Field Experiment in the Market for Electricity." *Proceedings of the National Academy of Sciences* 116, no. 12: 5293–5298. https://doi.org/10.1073/pnas.1802874115.

Brummett, C. M., R. Steiger, M. Englesbe, et al. 2019. "Effect of an Activated Charcoal Bag on Disposal of Unused Opioids After an Outpatient Surgical Procedure: A Randomized Clinical Trial." *JAMA Surgery* 154, no. 6: 558–561. https://doi.org/10.1001/jamasurg.2019.0155.

Buchmueller, T. C., and C. Carey. 2018. "The Effect of Prescription Drug Monitoring Programs on Opioid Utilization in Medicare." *American Economic Journal: Economic Policy* 10, no. 1: 77–112. https://doi.org/10.1257/pol.20160094.

Buttenheim, A., K. L. Milkman, A. L. Duckworth, D. M. Gromet, M. Patel, and G. Chapman. 2022. "Effects of Ownership Text Message Wording and Reminders on Receipt of an Influenza Vaccination: A Randomized Clinical Trial." *JAMA Network Open* 5, no. 2: e2143388. https://doi.org/10.1001/jamanetworkopen.2021.43388.

Buys, M., Z. Anderson, K. Bayless, et al. 2025. "Postsurgical Opioid Prescribing Among Veterans Using Community Care for Orthopedic Surgery at Non-Va Hospitals Compared to a Va Hospital With a Transitional Pain Service: A Retrospective Cohort Study." *Regional Anesthesia and Pain Medicine* 50, no. 7: 531–537. https://doi.org/10.1136/rapm-2023-105162.

CDC. 2020. Opioid Data Analysis and Resources.

CDC. 2023. Provisional Drug Overdose Death Counts.

Chater, N., and G. Loewenstein. 2023. "The i-frame and the S-Frame: How Focusing on Individual-Level Solutions Has Led Behavioral Public Policy Astray." *Behavioral and Brain Sciences* 46, no. e147: 1–84. https://doi.org/10.1017/s0140525x22002023.

Cutler, D. M., and J. T. Donahoe. 2024. *Thick Market Externalities and the Persistence of the Opioid Epidemic*. NBER Working Paper.

Cutler, D. M., and E. L. Glaeser. 2021. "When Innovation Goes Wrong: Technological Regress and the Opioid Epidemic." *Journal of Economic Perspectives* 35, no. 4: 171–196. https://doi.org/10.1257/jep.35.4.171. Accessed, Jan 1, 2024.

Dai, H., S. Saccardo, M. A. Han, et al. 2021. "Behavioural Nudges Increase Covid-19 Vaccinations." *Nature* 597, no. 7876: 404–409. https://doi.org/10.1038/s41586-021-03843-2.

Dimant, E., and S. Shalvi. 2022. "Meta-Nudging Honesty: Past, Present, and Future of the Research Frontier." *Current Opinion in Psychology* 47: 101426. https://doi.org/10.1016/j.copsyc.2022.101426.

Doctor, J. N., A. Nguyen, R. Lev, et al. 2018. "Opioid Prescribing Decreases After Learning of a Patient's Fatal Overdose." *Science* 361, no. 6402: 588–590. https://doi.org/10.1126/science.aat4595.

Dollar, S., H. Evans, J. Edwards, J. Odom-Forren, and B. Johnson. 2022. "Compliance With Opioid Disposal Following Opioid Disposal Education in Surgical Patients: A Systematic Review." *Journal of Peri-Anesthesia Nursing* 37, no. 4: 557–562. https://doi.org/10.1016/j.jopan.2021.10.017.

Duckworth, A. L., A. Ko, K. L. Milkman, et al. 2025. "A National Megastudy Shows That Email Nudges to Elementary School Teachers Boost Student Math Achievement, Particularly When Personalized." *Proceedings of the National Academy of Sciences* 122, no. 13: e2418616122. https://doi.org/10.1073/pnas.2418616122.

Eichmeyer, S., and J. Zhang. 2022. "Pathways into Opioid Dependence: Evidence From Practice Variation in Emergency Departments." *American Economic Journal: Applied Economics* 14, no. 4: 271–300. https://doi.org/10.1257/app.20210048.

Ericson, K. M. 2017. "On the Interaction of Memory and Procrastination: Implications for Reminders, Deadlines, and Empirical Estimation." *Journal of the European Economic Association* 15, no. 3: 692–719. https://doi.org/10.1093/jeea/jvw015.

Evans, W. N., E. M. Lieber, and P. Power. 2019. "How the Reformulation of Oxycontin Ignited the Heroin Epidemic." *Review of Economics and Statistics* 101, no. 1: 1–15. https://doi.org/10.1162/rest_a_00755.

FDA. 2023. Drug Disposal: Drug Take Back Locations.

Florence, C., F. Luo, and K. Rice. 2021. "The Economic Burden of Opioid Use Disorder and Fatal Opioid Overdose in the United States." *Drug and Alcohol Dependence* 218: 108350. https://doi.org/10.1016/j.drugalcdep.2020.108350.

Fryer, R. G. Jr, S. D. Levitt, J. List, and S. Sadoff. 2022. "Enhancing the Efficacy of Teacher Incentives Through Framing: A Field Experiment." *American Economic Journal: Economic Policy* 14, no. 4: 269–299. https://doi.org/10.1257/pol.20190287.

GAO. 2019. *Prescription Opioids: Patient Options for Safe and Effective Disposal of Unused Opioids*, Vol. 19–650, 1–24. GAO.

Gerber, A. S. and D. P. Green 2012. "Field Experiments: Design, Analysis, and Interpretation." Working Paper.

Gifford, E. D., K. T. Hanson, V. J. Davila, et al. 2020. "Patient and Institutional Factors Associated With Postoperative Opioid Prescribing After Common Vascular Procedures." *Journal of Vascular Surgery* 71, no. 4: 1347–1356. https://doi.org/10.1016/j.jvs.2019.05.068.

Gregorian, R., E. Marrett, V. Sivathanu, et al. 2020. "Safe Opioid Storage and Disposal: A Survey of Patient Beliefs and Practices." *Journal of Pain Research* 13: 987–995. https://doi.org/10.2147/jpr.s242825.

Hill, M., M. S. R. McMahon, and R. Barth. 2017. "Wide Variation and Excessive Dosage of Opioid Prescriptions for Common General Surgical Procedures." *Annals of Surgery* 265, no. 4: 709–714. https://doi.org/10.1097/sla.0000000000001993.

Howley, P., and N. Ocean. 2022. "Can Nudging Only Get You so Far? Testing for Nudge Combination Effects." *European Review of Agricultural Economics* 49, no. 5: 1086–1112. https://doi.org/10.1093/erae/jbab041.

Humphreys, K., C. L. Shover, C. M. Andrews, et al. 2022. "Responding to the Opioid Crisis in North America and Beyond: Recommendations of the Stanford–Lancet Commission." *Lancet* 399, no. 10324: 555–604. https://doi.org/10.1016/s0140-6736(21)02252-2.

Kahneman, D., J. L. Knetsch, and R. H. Thaler. 1991. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives* 5, no. 1: 193–206. https://doi.org/10.1257/jep.5.1.193.

Karlan, D., M. McConnell, S. Mullainathan, and J. Zinman. 2016. "Getting to the Top of Mind: How Reminders Increase Saving." *Management Science* 62, no. 12: 3393–3411. https://doi.org/10.1287/mnsc.2015.2296.

Lakens, D. 2022. Improving Your Statistical Inferences.

List, J. 2024. "Optimally Generate Policy-Based Evidence Before Scaling." *Nature* 626, no. 7999: 491–499. https://doi.org/10.1038/s41586-023-06972-y.

Liu, J. Y., J. C. Anderson, J. S. Franklin, F. A. Gesek, and D. I. Soybel. 2021. "Nudging Patients and Surgeons to Change Ambulatory Surgery Pain Management: Results From an Opioid Buyback Program." *Surgery* 170, no. 2: 485–492. https://doi.org/10.1016/j.surg.2021.01.016.

Liu, J. Y., J. S. Franklin, F. A. Gesek, and J. C. Anderson. 2020. "Buyback Program of Unused Prescription Opioids in Us Rural Communities, 2017–2018." *American Journal of Public Health* 110, no. 9: 1318–1324. https://doi.org/10.2105/ajph.2020.305730.

Lowenstein, M., J. Perrone, R. McFadden, et al. 2023. "Impact of Universal Screening and Automated Clinical Decision Support for the Treatment of Opioid Use Disorder in Emergency Departments: A Difference-In-Differences Analysis." *Annals of Emergency Medicine*. https://doi.org/10.1016/j.annemergmed.2023.03.033.

MacKinnon, J. G., M. Ø. Nielsen, and M. D. Webb. 2023. "Cluster-Robust Inference: A Guide to Empirical Practice." *Journal of Econometrics* 232, no. 2: 272–299. https://doi.org/10.1016/j.jeconom.2022.04.001.

Maclean, J. C., J. Mallatt, C. J. Ruhm, and K. Simon. 2021. "Economic Studies on the Opioid Crisis: Costs, Causes, and Policy Responses." In *Oxford Research Encyclopedia of Economics and Finance*.

Maclean, J. C., J. Mallatt, C. J. Ruhm, and K. Simon. 2022. "The Opioid Crisis, Health, Healthcare, and Crime: A Review of Quasi-Experimental Economic Studies." *Annals of the American Academy of Political and Social Science* 703, no. 1: 15–49. https://doi.org/10.1177/00027162221149285.

Maier, M., F. Bartoš, T. Stanley, D. R. Shanks, A. J. Harris, and E.-J. Wagenmakers. 2022. "No Evidence for Nudging After Adjusting for Publication Bias." *Proceedings of the National Academy of Sciences* 119, no. 31: e2200300119. https://doi.org/10.1073/pnas.2200300119.

Milkman, K. L., J. Beshears, J. J. Choi, D. Laibson, and B. C. Madrian. 2011. "Using Implementation Intentions Prompts to Enhance Influenza

Vaccination Rates.” *Proceedings of the National Academy of Sciences* 108, no. 26: 10415–10420. https://doi.org/10.1073/pnas.1103170108.

Milkman, K. L., L. Gandhi, M. S. Patel, et al. 2022. “A 680,000-Person Megastudy of Nudges to Encourage Vaccination in Pharmacies.” *Proceedings of the National Academy of Sciences* 119, no. 6: e2115126119. https://doi.org/10.1073/pnas.2115126119.

Milkman, K. L., D. Gromet, H. Ho, et al. 2021. “Megastudies Improve the Impact of Applied Behavioural Science.” *Nature* 600, no. 7889: 478–483. https://doi.org/10.1038/s41586-021-04128-4.

Milkman, K. L., M. S. Patel, L. Gandhi, et al. 2021. “A Megastudy of Text-Based Nudges Encouraging Patients to Get Vaccinated at an Upcoming Doctor’s Appointment.” *Proceedings of the National Academy of Sciences* 118, no. 20: e2101165118. https://doi.org/10.1073/pnas.2101165118.

Prendergast, M., D. Podus, J. Finney, L. Greenwell, and J. Roll. 2006. “Contingency Management for Treatment of Substance Use Disorders: A Meta-Analysis.” *Addiction* 101, no. 11: 1546–1560. https://doi.org/10.1111/j.1360-0443.2006.01581.x.

Pustejovsky, J. E., and E. Tipton. 2018. “Small-Sample Methods for Cluster-Robust Variance Estimation and Hypothesis Testing in Fixed Effects Models.” *Journal of Business & Economic Statistics* 36, no. 4: 672–683. https://doi.org/10.1080/07350015.2016.1247004.

Ruggeri, K., F. Stock, S. A. Haslam, et al. 2023. “A Synthesis of Evidence for Policy From Behavioural Science During Covid-19.” *Nature* 625, no. 7993: 1–14. https://doi.org/10.1038/s41586-023-06840-9.

Ruhm, C. J. 2019. “Drivers of the Fatal Drug Epidemic.” *Journal of Health Economics* 64: 25–42. https://doi.org/10.1016/j.jhealeco.2019.01.001.

Sacarny, A., E. Safran, M. Steffel, et al. 2022. “Effect of Pharmacist Email Alerts on Concurrent Prescribing of Opioids and Benzodiazepines by Prescribers and Primary Care Managers: A Randomized Clinical Trial.” *JAMA health forum, JAMA health forum*, American Medical Association, 3, no. 9: e223378. https://doi.org/10.1001/jamahealthforum.2022.3378.

Sacarny, A., D. Yokum, A. Finkelstein, and S. Agrawal. 2016. “Medicare Letters to Curb Overprescribing of Controlled Substances Had No Detectable Effect on Providers.” *Health Affairs* 35, no. 3: 471–479. https://doi.org/10.1377/hlthaff.2015.1025.

Soffin, E., L. Wilson, J. Liu, J. Poeran, and S. Memtsoudis. 2021. “Association Between Sex and Perioperative Opioid Prescribing for Total Joint Arthroplasty: A Retrospective Population-based Study.” *Drug and Alcohol Dependence* 126, no. 6: 1217–1225. https://doi.org/10.1016/j.bja.2020.12.046.

Substance Abuse and Mental Health Services Administration. 2024. “Key Substance Use and Mental Health Indicators in the United States: Results From the 2023 National Survey on Drug Use and Health.” Technical Report HHS Publication No. PEP24-07-021, NSDUH Series H-59.

Weiss, M. J., H. S. Bloom, and T. Brock. 2014. “A Conceptual Framework for Studying the Sources of Variation in Program Effects.” *Journal of Policy Analysis and Management* 33, no. 3: 778–808. https://doi.org/10.1002/pam.21760.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.
Supplorting Information S1: hec70085-sup-0001-suppl-data.pdf.