Incentives, Surrogates, and Long-run Vaccination

Pol Campos-Mercade Armando N. Meier Stephan Meier

Devin G. Pope Florian H. Schneider Erik Wengström*

June, 2026

Abstract

Can monetary incentives improve health behaviors in the long run, and do commonly used surrogate outcomes capture these effects? We study these questions in the context of vaccination using a large-scale field experiment. The experiment combines commonly used surrogates—vaccination intentions, intermediate behavioral proxies, and short-run vaccination—with long-run administrative vaccination records. We first document that incentives increase vaccination rates in the long run: guaranteed $20 incentives raise COVID-19 booster uptake by 9 percentage points. Lottery-based incentives also increase long-run uptake, while prosocial incentives primarily accelerate vaccination. Second, using surrogacy methods, we study whether surrogates can predict long-run impacts. Although the surrogates are strongly correlated with eventual vaccination, the assumptions required for surrogacy methods are often violated, and they do not accurately predict long-run impacts. Our findings highlight both the ability of incentives to change behavior and the importance of measuring long-run outcomes rather than relying solely on surrogates.

JEL Classifications: C93, D01, D62, I12, I18
Keywords: incentives, health behavior, vaccination, surrogates

* Pol Campos-Mercade, Department of Economics, Lund University, pol.campos@nek.lu.se; Armando N. Meier, Faculty of Business and Economics, University of Basel, armando.meier@unibas.ch; Stephan Meier, Columbia Business School, Columbia University, sm3087@gsb.columbia.edu; Devin Pope, Booth School of Business, University of Chicago, devin.pope@ChicagoBooth.edu; Florian H. Schneider, Department of Economics and Center for Economic Behavior and Inequality (CEBI), University of Copenhagen and CESifo, flsc@econ.ku.dk; Erik Wengström, Department of Economics, Lund University, erik.wengstrom@nek.lu.se

An earlier version of this paper circulated under the title "Incentives to Vaccinate." We are grateful to Marcella Alsan, Susan Athey, Josh Dean, Uri Gneezy, Alex Imas, Simon Reese, Pedro Rey-Biel, Claus Thustrup Kreiner, Avner Strulov-Shlain, Roberto Weber and seminar participants at Columbia University, Cornell University, Goethe University Frankfurt, Humboldt University, Erasmus University, Imperial College London, Lund University, Ludwig Maximilian University of Munich, Queen's University Belfast, University d'Aix-Marseille, University of Bologna, University of British Columbia, University of California San Diego, University of Chicago, University of Copenhagen, University of Essen, University of Lausanne, University of Manchester, University of Osaka, University of Pennsylvania (Wharton), University of Pittsburgh, University of St Andrews, University of Tokyo, and University of Verona for helpful comments and suggestions. We acknowledge funding from the Jan Wallanders och Tom Hedelius stiftelse samt Tore Browaldhs stiftelse, the Wallenberg Academy Foundation, the Swiss National Science Foundation Grant PZ00P1_201956, the Chazen Institute for Global Business at Columbia Business School, Columbia Business School, the University of Chicago Booth School of Business, and the Danish National Research Foundation Grant DNRF134.

1 Introduction

Whether monetary incentives can persistently change health behaviors is a fundamental question for policy-making. A large share of the public and private disease burden is linked to modifiable unhealthy behaviors (WHO, 2023), and economists and public health experts have long debated the potential of payments to promote healthy behavior and internalize externalities. For example, incentives have been considered to encourage exercising (e.g., Charness and Gneezy, 2009; Milkman et al., 2021*a*; Carrera et al., 2022), weight loss (Volpp et al., 2008), smoking cessation (Volpp et al., 2009), and blood donation (Lacetera, Macis and Slonim, 2013; Goette and Stutzer, 2020). A particularly critical application is incentives for vaccination. High vaccination rates are crucial for protecting public health and avoiding the economic costs of disease (Banerjee et al., 2010; Dai et al., 2021; Milkman et al., 2022*a*; Karing, 2024), but they are difficult to achieve. Incentives have therefore been frequently proposed to increase vaccination rates, although their use has been highly contested.[1]

A large literature finds that monetary incentives can impact stated intentions to vaccinate and short-run vaccination uptake (see discussion below). However, vaccination intentions and short-run uptake are typically used as proxy or "surrogate" outcomes, while the central policy question concerns long-run changes in vaccination rates. Importantly, it is unclear whether the documented effects on intentions and short-run uptake translate into persistent changes in vaccination rates. This reflects a broader pattern in empirical research: interventions are often evaluated using surrogate outcomes, which are cheaper and easier to collect, to draw conclusions about the policy-relevant long-run impacts. At the same time, a growing literature develops methods to infer long-run effects from surrogate outcomes (e.g., Battocchi et al. 2021; Athey, Chetty and Imbens 2025; Imbens et al. 2025). In this paper, we study whether monetary incentives affect long-run vaccination rates and, using surrogacy methods,

[1]This disagreement became especially visible during the COVID-19 pandemic; governments and organizations urgently sought ways to increase vaccination rates, and numerous academic and media articles debated the use of incentives. Some experts argued that incentives would likely be effective and that their benefits would outweigh potential risks (Savulescu, 2021; Oza, 2021; Persad et al., 2021). Proposals to pay people to take the vaccine were endorsed by journalists, policy-makers such as Joe Biden, and prominent researchers including Greg Mankiw, Steven Levitt, and Paul Romer. Others, however, cautioned against incentives, suggesting they might even reduce vaccination uptake (e.g., Jecker, 2021; Largent and Miller, 2021; Volpp, Loewenstein and Buttenheim, 2021). In practice, the absence of clear evidence on long-run effects and optimal design led to a variety of conflicting policy responses, ranging from lottery-based to guaranteed cash payments—or no incentives at all (for short-run analyses of some programs, see e.g. Barber and West 2022; Cohn et al. 2022; Milkman et al. 2022*b*; Thirumurthy et al. 2022).

assess whether these long-run effects can be reliably predicted from intentions and short-run vaccination behavior.

A key challenge in addressing both questions is data availability: credible causal inference requires exogenous variation in incentives, observation of both surrogate and long-run outcomes within the same sample, and sufficient statistical power. This paper overcomes these typical limitations by combining long-run administrative health records with short-run administrative and self-reported survey data from a large randomized controlled trial. This design allows us to make two main contributions. First, we provide evidence on whether monetary incentives generate persistent increases in vaccination rates or merely accelerate vaccinations that would have occurred anyway, by tracking outcomes over almost two years. Second, we evaluate whether the most commonly used short-run surrogates—such as stated intentions, intermediate behavioral proxies, and short-run effects—can reliably be used to predict these long-run treatment effects.

The experiment incentivized individuals in Sweden (N = 5,324) to receive a COVID-19 booster dose in 2022.[2] Participants were aged 18–64 and drawn from a broadly representative sample of the Swedish population in terms of age, gender, and region of residence. The Public Health Agency of Sweden recommended the vaccine to all individuals in this group, and the recruitment was timed to coincide with the national rollout of vaccinations. Individuals were randomly assigned to a control group or to one of four incentive conditions that provided a reward conditional on vaccinating within 30 days: a *guaranteed incentive* of about $20 (SEK 200), a *lottery* with an expected value of $20, a *donation* of $20 to charity,[3] or a *choice* condition allowing participants to select among these three options. We have access to long-run administrative records on all participants' vaccination dates over a 22-month horizon, enabling us to observe both short-run and long-run uptake. These data are combined with rich survey information on participants' socio-demographics and attitudes, as well as two commonly used short-run surrogates: stated vaccination intentions and clicking on a link to schedule a vaccination appointment.

[2]The setting closely resembles other vaccination settings—such as influenza, measles, or human papillomavirus—where individuals are encouraged to receive multiple doses over time and where vaccination uptake falls short of public health targets.

[3]These incentive types are among the most commonly used by academics and policy-makers (e.g., Mellström and Johannesson 2008; Lacetera and Macis 2010; Imas 2014; Tonin and Vlassopoulos 2015; DellaVigna and Pope 2018; Fabbri, Nicola Barbieri and Bigoni 2019; Campos-Mercade et al. 2021; Duch et al. 2023*a*).

We document two main findings. First, monetary incentives persistently increase vaccination rates. In the short run, all incentive schemes substantially raise vaccination rates, with guaranteed incentives being the most effective. In the control condition, 32% of participants receive the vaccine within 30 days. This share rises to 45% in the guaranteed incentive condition ($p < 0.001$), 41% in the lottery condition ($p < 0.001$), 37% in the donation condition ($p = 0.028$), and 43% in the choice condition ($p < 0.001$). The positive effects persist over time. After 22 months, vaccination uptake remains 8.6 percentage points higher in the guaranteed incentive condition ($p < 0.001$), 7.2 percentage points higher in the lottery condition ($p = 0.001$), and 6.8 percentage points higher in the choice condition ($p = 0.002$). In contrast, the long-run effect in the donation condition is only 0.6 percentage points higher and statistically indistinguishable from zero. Hence, for most treatments, the long-run effects are about two-thirds of the short-run effects, indicating that incentives accelerated vaccination among some individuals, but also persistently increased uptake among others. Using detailed data on socio-demographics and vaccination attitudes, we find that incentives increase vaccination rates across all subgroups that we can identify, including vaccine-hesitant individuals.[4]

Second, we find that relying on surrogate outcomes to draw conclusions about long-run effects leads to substantial mistakes. To examine this issue, we apply state-of-the-art surrogacy methods (Athey et al., 2025), developed to predict long-run treatment effects in contexts where only treatment effects on surrogate outcomes can be causally identified. Specifically, we assess whether long-run treatment effects can be predicted using survey measures (intentions and appointment-link clicks) and short-run vaccination behavior.

We find that intentions do not reliably predict either the direction or the magnitude of the long-run treatment effects. Contrary to actual behavior, intention data indicate null effects of most incentive conditions and even backfiring for the donation condition. When we examine appointment-link clicks as a surrogate, we find that while they correctly predict the direction of the effect, they do not adequately predict its magnitude. Our results show that relying on survey measures—such as intentions and appointment-link clicks—when providing policy guidance could be misleading and potentially lead to costly policy mistakes.

[4]In addition, we find no evidence that incentives reduce uptake of an additional COVID-19 booster vaccine (the fourth dose) one year later, consistent with Schneider et al. (2023). We also find that the incentives had positive spillover effects on the vaccination rates of participants' partners, yet no spillover effects on the vaccination rates of participants' children and parents.

These results may appear puzzling given the strong correlations we document between the survey-based surrogates and realized vaccination uptake. Correlations between survey measures and actual behavior are widely employed to validate the use of survey responses as proxies for behavior across the social sciences, including economics, psychology, public health, and related fields, and have contributed to a large literature relying on survey measures (see contributions below). However, a high correlation between survey responses and behavior does not guarantee that treatment effects on the survey measure translate into treatment effects on the behavioral outcome. If interventions alter the mapping between survey responses and realized behavior, as is the case in our study, validation based on correlations is insufficient to justify the use of surrogate outcomes for causal inference.[5]

The results for short-run vaccination rates are more nuanced. Vaccination uptake measured one month after participation provides a good approximation of the long-run treatment effects in our setting. However, other time horizons—both shorter windows of only a few weeks and longer windows of a few months—do not perform as well. Indeed, for these windows we can reject the key assumptions required to apply surrogacy methods. This pattern suggests that short-run uptake can offer a useful approximation of treatment effects, but its precision depends on the specific time horizon considered. Thus, although short-run uptake is substantially more informative than the other surrogates, its reliability as a surrogate cannot be taken for granted.

Our first contribution is to the literature on using monetary incentives to change health behaviors.[6] Part of this literature studies vaccination incentives. With notable exceptions among highly vaccine-hesitant populations (Jacobson et al., 2022; Chang et al., 2023), most experimental studies document positive effects of guaranteed incentives on vaccination intentions and short-run vaccination uptake across a range of countries and vaccine contexts (Bronchetti, Huffman and Magenheim 2015; Alsan, Garrick and Graziani 2019; Campos-Mercade et al. 2021; Duch et al. 2023*b*; Shen et al. 2024; Chang et al. 2026; see Campos-

[5] In our setting, participants who report an intention to vaccinate are substantially more likely to follow through and get vaccinated during the eligible period when offered incentives, illustrating that the intention–behavior relationship itself is treatment-dependent.

[6] Examples include studies on encouraging exercising (Charness and Gneezy, 2009; Royer, Stehr and Sydnor, 2015; Carrera et al., 2020; Milkman et al., 2021*a*), smoking cessation (Volpp et al., 2006, 2009), weight loss (Volpp et al., 2008), blood donation (Mellström and Johannesson, 2008; Lacetera, Macis and Slonim, 2013, 2014; Goette and Stutzer, 2020), healthy eating (List and Samek, 2015; Dolan, Galizzi and Navarro-Martinez, 2015; Belot, James and Nolen, 2016; Angelucci et al., 2019), health screening (Lieberman et al., 2019; Jones, Molitor and Reif, 2024), sleeping (Giuntella, Saccardo and Sadoff, 2024), organ donation (Becker and Elias, 2007), and medication adherence (Barankay et al., 2020).

Mercade et al. 2023 for a review).[7,8] Campos-Mercade et al. (2021) is the closest study to the present paper. It examines the short-run effects (30 days) of a guaranteed SEK 200 incentive on first-dose COVID-19 vaccination using a broadly representative Swedish sample linked to nationwide administrative data. Schneider et al. (2023) use the same sample to study potential unintended consequences of monetary incentives on subsequent unincentivized doses, as well as on attitudes and other behaviors, and find no evidence of crowding out.

However, because prior studies observe either survey-based intentions or short administrative windows, they cannot determine whether these positive effects reflect persistent increases in vaccination or simply acceleration. We address this gap by providing experimental evidence on the long-run effects of monetary incentives on vaccination. By combining randomized assignment with comprehensive individual-level administrative vaccination records spanning nearly two years, we distinguish between short-run acceleration and long-run changes in vaccination.

Our second contribution is to the literature on surrogate outcomes. In many empirical and policy settings, researchers and policy-makers observe self-reported survey measures or short-run outcomes—such as stated intentions, early behavioral responses, or short-run administrative data—but lack access to the long-run outcomes of interest. There are two main approaches to addressing this issue. First, much of the literature treats a surrogate outcome as a valid proxy for the outcome of interest whenever existing evidence documents a correlation between the surrogate and the outcome (e.g., Dohmen et al., 2011; Falk et al., 2018, 2023; Wiswall and Zafar, 2018; Cavatorta and Schröder, 2019; Enke, Rodriguez-Padilla and Zimmermann, 2022; Buser, Niederle and Oosterbeek, 2024). This approach implicitly assumes that such a correlation implies that the treatment effects on the surrogate will be comparable to the treatment effects on the outcome of interest.[9] At the same time, a rapidly

[7]The largest body of research examines the impact of incentives on vaccination intentions (e.g., Klüver et al. 2021; Serra-Garcia and Szech 2023; Ruggeri et al. 2024; for meta-studies with further references, see Batteux et al., 2022; Huang, Huang and Yu, 2023; Khazanov et al., 2023). The few studies that observe actual behavior focus only on short-run effects. For example, they study whether incentives increase flu shots taken on the same day or week (Alsan, Garrick and Graziani, 2019; Chang et al., 2026), or COVID-19 vaccination over a few weeks (Campos-Mercade et al., 2021; Milkman et al., 2022*b*; Chang et al., 2023).

[8]The results on the effects of the US state incentive lottery programs, which offered large prizes with very small probabilities, on vaccination uptake are also mixed. While Barber and West (2022) and Cohn et al. (2022) report positive effects, Milkman et al. (2022*b*) and Thirumurthy et al. (2022) found minimal effects, if any, of various city and state incentive programs.

[9]This premise underlies a vast empirical literature on vaccination which evaluates interventions using stated intentions as the outcome variable (e.g., Betsch et al. 2017; Lau et al. 2019; Ashworth et al. 2021; Klüver et al. 2021; Steinert et al. 2022; Angerer et al. 2023; Serra-Garcia and Szech 2023; Moehring et al.

growing and independent literature formalizes and develops econometric methods that use correlations between surrogate and long-run outcomes to predict the magnitude of long-run treatment effects when only the effects on the surrogate outcomes are causally identified (e.g., Prentice 1989; VanderWeele 2013; Anderer, Bastani and Silberholz 2021; Battocchi et al. 2021; Obradović 2024; Yang et al. 2024*a*; Athey, Chetty and Imbens 2025; Athey et al. 2025; Imbens et al. 2025).

Our setting allows a direct assessment of these approaches, as we observe treatment effects on multiple short-run surrogate outcomes and on long-run vaccination behavior within the same experimental sample.[10] We also illustrate that insights from behavioral economics—such as the idea that interventions can affect the likelihood that individuals follow through on their intentions (e.g., Milkman et al. 2011)—help explain why the assumptions required to use surrogate outcomes to predict long-run treatment effects may fail.

Overall, our findings have broader implications for research that relies on surrogate outcomes. Collecting long-run administrative or behavioral outcomes is often difficult, costly, and slow, leading many literatures to rely mainly on survey measures or short-run proxies that are easier to obtain.[11] Our results suggest that such reliance can be misleading even when these surrogates are correlated with the ultimate outcomes of interest. In our setting, commonly used survey-based measures and intermediate behavioral proxies fail to predict the direction or magnitude of long-run treatment effects, despite being strongly correlated with

2023; Alsan and Eichmeyer 2024). Systematic reviews document several hundred such studies across contexts and populations (see, e.g., Jarrett et al. 2015; Brewer et al. 2017; Batteux et al. 2022; Malik et al. 2023).

[10]Related papers compare treatment effects on surrogate outcomes and actual behaviors of interest measured in different populations. For example, Dai et al. (2021) and Saccardo et al. (2024) compare treatment effects on intentions among online samples with actual treatment effects on behavior among the general population. These studies make the important point that results based on surrogate outcomes from convenience samples—which are often used due to low costs—cannot be directly generalized to actual field behaviors. However, such comparisons do not allow validating surrogacy methods; it is not possible to determine whether any differences are driven by population differences or by limitations of the surrogate measures. More formally, our design satisfies the comparability condition within the experimental sample (Athey et al., 2025), allowing a direct assessment of surrogate validity.

[11]Studies in education, labor, and organizational economics, for example, often focus on self-reported outcomes such as motivation, intended college attendance, intended major choice, preferences over job characteristics elicited in hypothetical vignettes, or behavioral proxies such as clicking a job listing or requesting more information about a vacancy, signing up for job alerts, submitting a CV, signing up for SAT/ACT prep resources, job search behaviors or short-run employment, while the ultimate outcomes of interest may be long-run measures such as actual career choices or lifetime earnings. Similarly, much work in macroeconomics and finance relies on spending plans and saving intentions, and research in public economics and political economy often studies voting intentions, small donations to political parties, clicking a link to voter registration or signing up for campaign newsletters (see, for example, many of the papers discussed in Fuster and Zafar 2023; Haaland, Roth and Wohlfart 2023; Stantcheva 2023).

the long-run outcomes.[12] These findings indicate that collecting longer-run administrative or behavioral data may be substantially more valuable than is often assumed, particularly in contexts where surrogate outcomes have little intrinsic value absent persistent behavioral change. Moreover, they highlight the importance of carefully evaluating the assumptions required to extrapolate from surrogate outcomes and suggest that researchers and policy-makers should be cautious when drawing conclusions about long-run behavioral impacts from short-run or survey-based measures alone.

This paper proceeds as follows. In Section 2, we document the experimental design and data. Section 3 presents our findings on the effects of incentives on short-run and long-run vaccination. Section 4 analyzes the extent to which long-run effects can be predicted from short-run surrogates. Section 5 studies cost-effectiveness. Section 6 concludes.

2 Experimental design, data and estimation

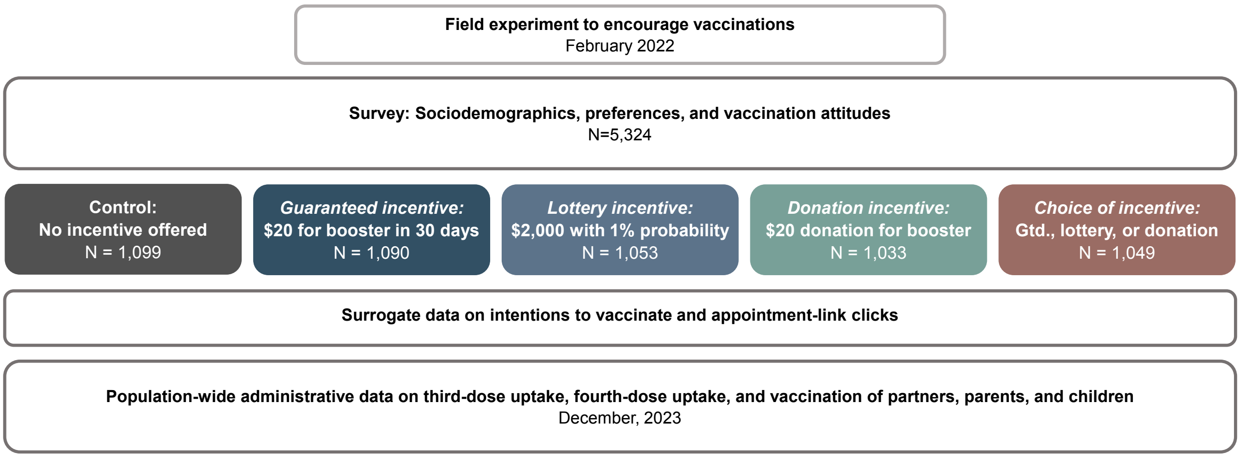
We conducted a pre-registered randomized controlled trial (RCT) in a general population sample of Swedish residents to study the effects of monetary incentives on vaccination. The experiment combines randomized assignment with survey data and population-wide administrative vaccination records, enabling us to track vaccination behavior over time. The survey data were linked to administrative records by the Public Health Agency of Sweden and the Swedish Tax Agency. The study protocols were approved by the Swedish Ethical Review Authority (DNR 2021-06669-01).

2.1 Experiment overview

Figure 1 summarizes the experimental design. We conducted the field experiment with 5,324 participants aged 18 to 64, recruited from a broadly representative online panel by the survey company Norstat. These age groups had been recommended to receive two doses of a COVID-19 vaccine in 2021. At the time of the trial, individuals who had received both doses were recommended to receive a booster dose (the third dose). We carried out the experiment right as booster vaccination became available for these age groups, in February 2022. For the age groups of our participants, only the Pfizer-BioNTech and Moderna vaccines were used,

[12]This insight is consistent with Chapman et al. (2025), who show that validation based on average correlations can mask important heterogeneity across groups.

Figure 1: Experimental design



Note: The figure summarizes the experimental design with the corresponding treatment groups and the linked experimental, survey, and population-wide administrative data. Using the participants' social security number, the Public Health Agency of Sweden matched the survey data with population-wide Swedish administrative vaccination records, including records for family members. As it is not possible to opt out of or delete records in the vaccination registry, the administrative records include the date of each COVID-19 vaccination of each Swedish resident.

which were offered free of charge. Individuals could easily sign up online or by phone for a vaccination appointment at their local vaccination center with typically no waiting time.[13]

The Swedish government did not implement any restrictions for people who did not get a booster dose. Accordingly, our experiment captures the effects of incentives for vaccination doses which are free, do not affect individual movement restrictions, and are targeting a disease that is essentially endemic. These conditions are similar to those for other important vaccines, such as influenza, measles, tuberculosis, human papillomavirus (HPV), and potential malaria vaccines.

In a pre-screening questionnaire administered by Norstat, we restricted participation to individuals who had received the two initial COVID-19 vaccine doses and were therefore eligible and recommended to receive a booster.[14] There had been substantial pressure to receive the initial two COVID-19 vaccine doses, including travel restrictions, and by the time of the study, approximately 90% of the Swedish population had received them (OWID, 2022).

[13] While the initial distribution of the first dose varied regionally, by July 2021, the vaccine was accessible to all adults. Individuals were advised to receive their second dose 3-4 weeks following the first. Additionally, a booster was recommended at least 90 days after the second dose. Consequently, everyone in the age groups we examined had the opportunity to receive their second dose well before getting the booster in February 2022.

[14] A small number of individuals enrolled even though they had not yet received two doses or had already received a booster dose. Our analysis excludes these participants, but all results are robust to including them (Appendix Tables B.6 and B.7).

As a result, even among vaccinated individuals, there remained substantial heterogeneity in vaccination attitudes (see discussion in Section 3.3).

After the pre-screening and consent form, we gathered participants' social security number (personnummer) and measured their socio-demographics, economic preferences, personality traits, and attitudes towards COVID-19 vaccines. Participants were then randomly assigned to either a control group or one of several incentive conditions, each offering a reward conditional on receiving a booster dose within 30 days. At the end of the survey, participants reported their intention to vaccinate and were given the opportunity to click a link to book a vaccination appointment.[15]

Using the participants' social security number, the Public Health Agency of Sweden matched the survey data with population-wide Swedish administrative records for vaccinations, which allow us to examine whether and when each participant got vaccinated. As the Public Health Agency matched the data in December 2023, we can examine long-run effects of incentives up to 660 days after RCT participation. The data include the entire COVID-19 vaccination history, including fourth-dose uptake, allowing us to study the long-run effects of incentives on future vaccination uptake once monetary incentives were no longer offered. Providing social security numbers is common in Sweden, and we explained in the consent form that, even after linkage to other data, the social security number would remain completely anonymous to us as researchers.[16]

[15] The full questionnaire, translated into English, is available on OSF: https://osf.io/ra3mv/.

[16] The first two screens of the online experiment were hosted by the survey provider and asked participants about their social security number. The participants were then forwarded to the actual experiment hosted by us, with a newly constructed ID. We did not have access to the dataset containing the social security numbers, and the survey company did not have access to our dataset. Instead, the survey provider directly sent the dataset consisting of the social security number and the constructed ID to the Public Health Agency of Sweden. We then sent the experimental data that included the constructed ID to the Public Health Agency. The Public Health Agency then linked the two datasets, linked this data to registry data, and then deleted the social security number and the newly constructed ID. We then received a completely anonymized dataset. The reason why we opted for this procedure was to guarantee that the social security number would never be stored together with the experimental data or the vaccination data, and to increase participants' trust when providing their social security number. Note that people in Sweden are used to providing their social security number on many occasions, including most online transactions and memberships.

2.2 Survey

Questionnaire. After providing consent and prior to treatment assignment, participants completed a survey. We first elicited a set of economic preferences.[17] We then elicited attitudes toward the COVID-19 vaccine to capture vaccine hesitancy (concerns about side effects, perceived safety, and knowledge about side effects), information on COVID-19 history (prior infection, risk group status, and concerns about infection), and standard socio-demographic characteristics.

Experimental conditions. After completing the questionnaire, participants were randomly allocated to one of five experimental conditions. In the control condition, as in all other experimental conditions, we encouraged participants to vaccinate within 30 days of participating in the survey ("We would like to encourage you to get the third dose of a COVID-19 vaccine as soon as it becomes available to you, ideally within 30 days of taking part in this survey").

In addition to the message that was also part of the control condition, the four incentive conditions included rewards conditional on vaccination within 30 days after participating in the survey. In the guaranteed incentive condition, we offered participants the equivalent of \$20 in Swedish kronor (SEK 200). In the lottery incentive condition, we offered participants a lottery ticket with a 1% chance of winning \$2000 (the expected value is \$20). In the donation incentive condition, we offered to make a donation of \$20 to the charity Save the Children on the participant's behalf. Finally, in the choice condition, we offered people a free choice among the three incentive schemes (guaranteed, lottery, or donation). For example, if a participant chose the guaranteed incentive, she would receive \$20 conditional on vaccination within the 30-day window.[18]

We informed participants that we would verify their vaccination status and timing using administrative data and provide the promised reward if they got vaccinated within the 30-

[17] We measured altruism, reciprocity, trust, patience, and risk attitudes using the survey instruments from the Global Preference Survey (Falk et al., 2018). In addition, we collected measures of procrastination, self- and social-image concerns in the prosocial domain, and extrinsic motivation (Tuckman, 1991; Amabile et al., 1994; Plant and Devine, 1998; Aquino and Reed II, 2002), as well as adherence to social norms and the importance of autonomy using items from the Schwartz Value Survey. We then included incentivized questions to elicit prosociality and risk preferences using a dictator game and an investment game (see Appendix Section D.1 for details).

[18] In the choice condition, 42.3% of participants selected the guaranteed payment, 28.0% chose the lottery payment, and 29.6% opted for the donation. Choices were strongly predicted by measured preferences: more prosocial individuals were more likely to select the donation, and more risk-tolerant individuals were more likely to select the lottery. See Appendix Section D.3 for details.

day window.[19] Additionally, we included a small text box in the survey where one of the authors signed a statement that participants would receive the reward if they were vaccinated within the specified time frame.[20] Participants were also informed that we would match the experimental data with vaccination registries to facilitate the payments as soon as possible, but no later than July 2022. The payment was made using "Norstat cash," a system familiar to participants that allows them to redeem their earnings for vouchers valid at most major online and in-person retail stores in Sweden.

Surrogate survey outcomes: Intentions and clicking on an appointment link. Following treatment assignment, we measured two commonly used short-run proxy outcomes. First, we elicited participants' vaccination intentions by asking whether they intended to receive a booster dose of a COVID-19 vaccine within one month (*Intention Vaccinate*). Second, at the end of the survey, we provided a link to the website of the relevant regional health authority containing information on how to book a vaccination appointment. We recorded whether participants clicked on this link (*Appointment Link Click*). Because both outcomes were measured after treatment assignment, they may reflect immediate behavioral or attitudinal responses to the incentive conditions and thus serve as potential surrogates for realized vaccination uptake.

[19]Because the third dose was only recommended and available three months after the second dose, in all conditions, including the control group, we clarified that "if it has not yet been three months since you took the second dose, then we mean 30 days after the third dose becomes available to you." In the incentive conditions, we told participants that if it had been less than three months since the second dose, to pay the incentive we would instead consider the 30-day window after the third dose became available to them. For most participants this corresponds to the 30-day window after the trial. However, for 415 participants it had been less than three months since they received the second dose when they participated in the trial. Because the third dose was only recommended and available three months after the second dose, the 30-day incentive window for these participants started after the trial. For these participants, we consider the 30-day window that started three months after they got the second dose. This is a slight deviation from the pre-registration plan, where we made the mistake of considering the 30-day window after the trial for all participants, which was inconsistent with the experimental design. However, in Appendix Table B.5, we show that the results do not change if we instead follow this pre-registered approach or if we exclude participants who received the second dose less than three months before the study.

[20]One potential concern is that the share of participants that finish the survey differs across treatment conditions, for example, because the different conditions may imply different intellectual effort. However, 99% of the participants who encountered the treatments finished the survey and we find no meaningful differences across the treatments, indicating no meaningful treatment differences in sample selection. An advantage is that we asked for the social security number at the very beginning of the study. Hence, we can also match to administrative data all participants who dropped out when they encountered the treatments. In our main specification, we include all participants who encountered one of the conditions, even if they did not finish the study after that, to avoid any potential bias.

2.3 Administrative vaccination records

We linked the experimental data with administrative data from national COVID-19 vaccination registers comprising all residents of Sweden. As it is not possible to opt out of or delete records in the vaccination registry, the administrative records include the date of each COVID-19 vaccination of each Swedish resident.

In the administrative data, we see whether and when each participant got vaccinated. The Public Health Agency of Sweden linked our trial data at the individual level with the administrative data on December 19, 2023. As the trial ended on February 8, 2022, we observe for each participant whether and when they got vaccinated within more than 22 months after trial participation. Our pre-registered main outcome variable corresponds to a dummy variable on whether participants vaccinated within the 30-day window (Vaccination uptake within 30 days). We similarly construct vaccination uptake within 15, 60, 90, 120, 150, 360, and 660 days after the 30-day window started. The data also allow us to measure fourth-dose uptake up to December 2023.

Finally, in collaboration with the Swedish Tax Agency, we link participants to complete administrative vaccination records of their registered partners, parents, and children. This registry linkage allows us to quantify spillover effects within families (see Appendix Section D.5 for details).

2.4 Data collection, exclusion criteria, and sample

The participants were recruited from a general population panel in Sweden by the survey company Norstat. Norstat actively recruits people mainly by means of phone calls to create a panel that is broadly representative in terms of age, gender, and region, and compensates participants for answering a few surveys each year. In our case, we asked the company to recruit participants between the ages of 18 and 64. Participants were asked to fill out an online survey, and responses were collected between January 27, 2022 and February 8, 2022.

Participants were ineligible for the trial if they had not yet received two doses of a COVID-19 vaccine, already received a booster dose, or were not recommended to take a booster dose by the Public Health Agency of Sweden at the time of the trial. Eligibility was assessed in the consent form, where participants were asked whether they had not yet received two doses of a COVID-19 vaccine, had already received a booster dose, were pregnant, had previously experienced an allergic reaction that required hospital care, or had ever experienced a severe allergic reaction after vaccination. Participants who answered affirmatively to any of these questions were screened out and could not proceed to the experiment.

We received 6,579 responses to the online survey. Based on administrative records, we first exclude individuals who did not satisfy eligibility for the study: 612 who had already received a booster dose, 232 who had not yet received a second dose, and 19 who were younger than 18 at the time of the trial. We further exclude 197 individuals who discontinued the survey before reaching the treatment or control screen, and 195 duplicate responses from individuals who completed the survey more than once and were assigned to different treatment conditions. We retain all other observations, including 69 participants who were exposed to the intervention but did not complete the survey, to avoid post-treatment selection bias. The final analysis sample consists of 5,324 participants. Appendix Tables B.6 and B.7 show that the results are robust to alternative sample inclusion criteria.

Descriptive statistics of the main trial sample are presented in Appendix Tables A.1 and A.2. As in any survey-based sample, unobserved characteristics may differ from the broader population. However, along observable dimensions, our sample is broadly representative of the Swedish population in terms of gender, region, income, education, and vaccination rate. We do, however, have an overrepresentation of individuals aged 26 to 35 years and an underrepresentation of individuals aged 46 and older (see Appendix Table A.2). In Appendix Table B.1, we show that results do not change when using sampling weights to adjust for the misrepresentation. In addition, we find that participants' socio-demographics are comparable across experimental conditions (see Appendix Table A.3).

2.5 Estimation and pre-registration

We pre-registered our analyses in detail in the AEA RCT Registry (AEARCTR-0008906). To estimate treatment effects on vaccination uptake, we pre-registered using ordinary least squares (OLS) regressions exactly as follows:

$$\begin{aligned}\text{Vaccination}_i = \beta_0 &+ \beta_1 \mathbb{1}(\text{``guaranteed''})_i + \beta_2 \mathbb{1}(\text{``lottery''})_i + \beta_3 \mathbb{1}(\text{``donation''})_i \\ &+ \beta_4 \mathbb{1}(\text{``choice''})_i + \beta_5 X_i + \epsilon_i\end{aligned}$$

where Vaccination_i is the outcome of interest for participant i and $\mathbb{1}(\text{"treatment"})_i$ is a dummy variable capturing whether the participant was assigned to the specified treatment condition. The impact of each treatment relative to the control condition is captured by the corresponding β coefficient. X_i is a vector of pre-registered control variables consisting of gender, age groups in 5-year brackets, region, interactions between each age group and region, being in an at-risk group for COVID-19, civil status, having children in the household, employment status, education, parents' place of birth, and income. Finally, ϵ_i is an individual-specific error term, and we report heteroskedasticity-robust standard errors.

Our pre-registration plan focused on estimating the causal effects of the incentive conditions on realized vaccination uptake. At the time of registration, we specified vaccination within 30 days as the primary outcome, as it was uncertain how long administrative follow-up would be available. In the paper, we report effects both at the pre-registered 30-day horizon and over the maximum available follow-up period (660 days), which allows us to study long-run impacts and distinguish between acceleration and durable increases in vaccination. The empirical specifications follow the pre-registered analysis plan. While intentions and link clicks were pre-registered as outcomes of interest, the use of formal surrogate outcome methods to assess their predictive performance was not pre-specified.

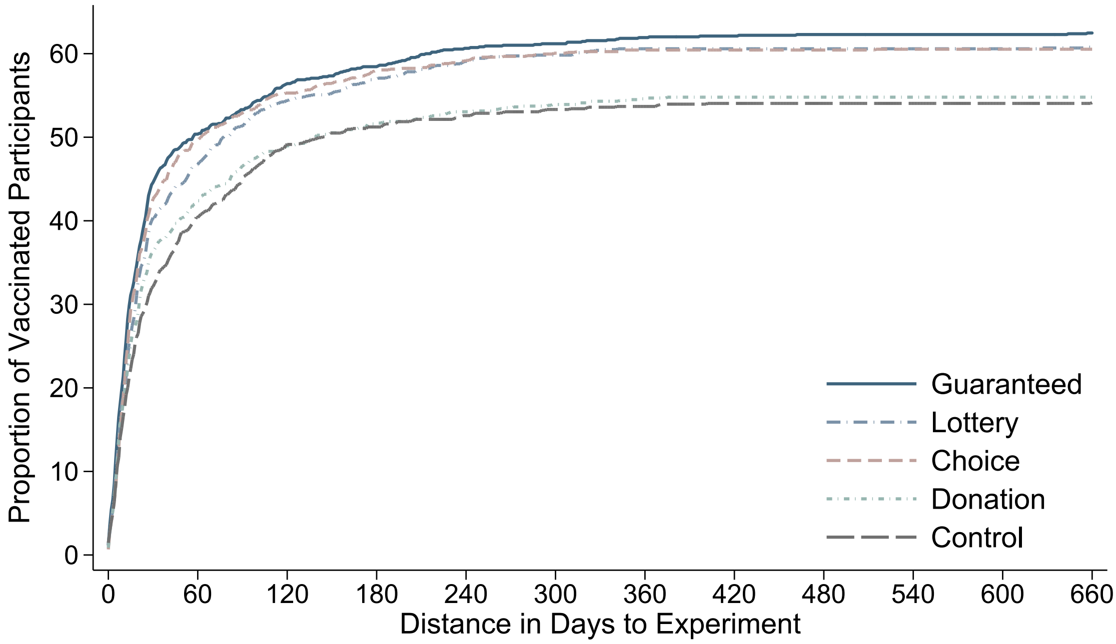
3 Treatment effects of incentives for vaccination

This section studies the effects of incentives on vaccination uptake. We first document treatment effects on vaccination in the short run (within 30 days of study participation) and in the long run (within 660 days of study participation). We then present analyses of heterogeneous treatment effects, to assess potential backfiring for subgroups, along with additional results on subsequent dose uptake and spillovers.

3.1 Effects of incentives on short-run vaccination uptake

Figure 2 plots vaccination uptake over time relative to trial participation for all treatment groups (Kaplan–Meier curves). The figure shows that guaranteed incentives generate the largest increase in short-run vaccination uptake, followed by the choice condition and the lottery incentive. In the guaranteed condition, vaccination within 30 days increases from

Figure 2: Incentives increase vaccination uptake



Note: The figure shows the development of the proportion of participants who got vaccinated over time in the different treatment conditions (Kaplan-Meier curves). The figure is based on data from the trial linked to Swedish administrative vaccination records, which include the date each participant got vaccinated ($N = 5,324$).

32.3% in the control group to 45.2%, a 12.9 percentage points (pp) increase corresponding to a relative increase of about 40%.

Table 1, column (2), shows the corresponding regression estimates, using the pre-registered specification and controls. The regression results confirm the pattern in the raw data: guaranteed incentives increased vaccination uptake within 30 days by 13.02 pp ($p < 0.001$), lottery incentives by 8.45 pp ($p < 0.001$), the donation incentive by 4.66 pp ($p = 0.028$), and the choice condition by 10.90 pp ($p < 0.001$).

Notably, the treatment effect of guaranteed incentives is substantially larger than what we found for similar incentives for the first dose in Campos-Mercade et al. (2021). In that study, we report that such incentives increased first-dose uptake from 71.6% in the control condition to 75.8% in the incentive condition. We believe that differences in the policy environment help explain this discrepancy. During the initial vaccination rollout, there was substantial public pressure to receive the first two COVID-19 vaccine doses. As a result, vaccination uptake was already very high in the control condition, limiting the scope for incentives to further increase vaccination. In contrast, the booster-shot setting is closer to the typical environment for recurrent vaccinations, where vaccination is recommended but

Table 1: Impact on short- and long-run vaccination uptake

Dependent Variable	Vaccination uptake within days							
	15	30	60	90	120	150	360	660
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Guaranteed Condition	9.94***	13.02***	9.95***	8.25***	7.03***	6.66***	8.44***	8.58***
	(1.94)	(2.11)	(2.16)	(2.17)	(2.17)	(2.17)	(2.14)	(2.14)
Lottery Condition	4.06**	8.45***	6.45***	6.73***	5.27**	4.87**	7.26***	7.18***
	(1.90)	(2.12)	(2.17)	(2.18)	(2.18)	(2.17)	(2.14)	(2.14)
Donation Condition	3.21*	4.66**	1.84	1.66	-0.09	0.24	0.84	0.61
	(1.92)	(2.12)	(2.19)	(2.21)	(2.21)	(2.21)	(2.20)	(2.20)
Choice Condition	6.71***	10.90***	9.01***	8.48***	6.11***	5.94***	7.16***	6.83***
	(1.95)	(2.13)	(2.18)	(2.20)	(2.20)	(2.19)	(2.16)	(2.16)
Pre-reg. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Control Cond.	22.65	32.30	40.76	45.04	49.31	50.68	53.69	54.14
Observations	5,324	5,324	5,324	5,324	5,324	5,324	5,324	5,324

Note: The table shows coefficient estimates from linear regressions of short- and long-run vaccination uptake on indicators for the experimental conditions using the pre-registered set of controls. "Vaccination uptake within X days" measures the proportion of participants who got vaccinated within X days after the start of the incentive window. Heteroskedasticity-robust standard errors are shown in parentheses. "Mean Control Cond." gives vaccination uptake after X days in the control condition. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

not accompanied by comparable mandates or pressure. In such more standard settings, there may be substantially greater scope for interventions to increase vaccination uptake.

While the pre-registered main analysis focuses on comparisons between each incentive condition and the control condition, we can also compare the incentive conditions with one another. The guaranteed condition had the largest effect. Vaccination uptake after 30 days in the guaranteed condition is statistically significantly higher than in the lottery ($p = 0.034$) and donation conditions ($p < 0.001$), but not higher than in the choice condition ($p = 0.327$).[21]

Hence, we find that all conditions were successful in increasing short-run vaccination rates. However, the key policy question is whether these differences persist in the long run or merely reflect acceleration.

[21] We also find that vaccination uptake is higher in the choice condition than in the donation condition ($p = 0.004$). The differences between the lottery and the choice condition ($p = 0.260$) and between the lottery and the donation condition ($p = 0.082$) are not statistically significant.

3.2 Effects of incentives on long-run vaccination uptake

Table 1, in columns (3) to (8), provides results on intermediate and long-run vaccination rates. We find that incentives not only accelerated vaccination uptake but increased vaccination rates in the long run. While the effect sizes get smaller in absolute terms—implying that part of the increase is due to an acceleration of vaccine uptake—they remain large and statistically significant for the guaranteed, lottery, and choice conditions. Even after 22 months, uptake remains 8.58 pp higher in the guaranteed condition ($p < 0.001$), 7.18 pp higher in the lottery condition ($p = 0.001$) and 6.83 pp higher in the choice condition ($p = 0.002$). In the donation condition, the estimated treatment effect after 22 months is only 0.61 pp, which is not statistically different from zero ($p = 0.781$).

Figure 2 also shows persistently higher vaccination rates in the guaranteed, lottery, and choice conditions, but not in the donation condition. There is an initial increase in vaccination rates in the donation condition, but the control condition catches up within 2 months after trial participation. For the other treatments, the initial increases translate into persistent differences in vaccination rates, with the guaranteed condition dominating the other conditions throughout.[22]

Taken together, these results show that guaranteed incentives are particularly effective in both accelerating and persistently increasing vaccination uptake. The lottery and the choice conditions are also effective, particularly in the long run, while the donation condition is less effective in the short run and does not lead to a persistent increase in vaccination rates in the long run. These results are robust to a battery of robustness checks, such as including different sets of control variables, using different sample inclusion criteria, using sample weights, and using logistic regressions (Appendix Section B).

3.3 Heterogeneity and vaccination uptake

To assess the generalizability of our results, we draw on the extensive data on vaccine attitudes and socio-demographic characteristics collected for all participants to study heterogeneous treatment effects (see results in Appendix Section D.1 and D.2). Although our sample includes only individuals who previously received COVID-19 vaccine doses, it also contains

[22] We find consistent treatment effects on days to vaccinate using a Tobit regression and on hazard ratios using a Cox proportional hazards model (Appendix Table B.2).

a substantial share of participants with negative vaccine attitudes.[23] We find little evidence of heterogeneous treatment effects by vaccine hesitancy, although the estimates suggest that interventions are somewhat less effective among more hesitant individuals. In almost all specifications, even the most vaccine-hesitant groups exhibit positive treatment effects.[24]

Across all other captured dimensions—including prosociality, risk preferences, COVID-19 history, income, age, and education—we do not detect meaningful heterogeneous treatment effects. Confidence intervals are generally sufficiently tight to rule out economically relevant negative effects. Overall, we find that incentives tend to increase vaccination rates across all subgroups that we can identify, with no evidence of backfiring in any subgroup. These findings also speak to concerns in the literature that financial incentives may backfire for certain groups (e.g., Titmuss 1970; Sandel 2012; Gneezy, Meier and Rey-Biel 2011; Bowles and Polanía-Reyes 2012).

3.4 Additional results

Beyond the direct effects on vaccination uptake, our data allow us to examine (i) subsequent, unincentivized vaccination decisions and (ii) spillover effects on family members.

Nearly two years after the intervention, 9.5% of control-group participants received a fourth dose. We find no statistically or economically significant impacts of any incentive condition on fourth-dose uptake after 660 days of trial participation (see Appendix Section D.4). The estimated treatment effects in the guaranteed, lottery, donation, and choice conditions are 0.65 pp ($p = 0.592$), 0.02 pp ($p = 0.988$), 0.04 pp ($p = 0.977$), and -0.57 pp ($p = 0.630$), respectively. Pooled across incentive arms, the estimated effect is 0.04 pp ($p = 0.96$, 95% CI: $[-1.82, 1.91]$). While the estimates are somewhat imprecise, they do not provide evidence of

[23] One indication is that uptake of the booster dose in the control group was only about 50%, despite the clear recommendation for everyone to receive it. Moreover, 23.6% do not agree with the statement that COVID-19 vaccines are safe, and 35.7% do not disagree with the statement that vaccination can trigger diseases. These individuals may have received the initial doses but subsequently become more skeptical, or they may have been vaccinated earlier due to regulations such as travel or work requirements.

[24] To assess how including individuals who had not received the first two vaccine doses would affect our results, we conduct a simple calibration exercise. While the data suggest that incentives also increase vaccination among hesitant individuals, we take a conservative approach and assume zero treatment effects for the previously unvaccinated. (Note that incentives are very unlikely to backfire for this group, as this population would not vaccinate in the control condition either.) As around 10% of the population had not received the first two doses at the time of the study, the implied long-run effect for the guaranteed condition becomes $0.9 \times 8.58 + 0.1 \times 0 = 7.72$. The intervention would thus remain highly effective even under this conservative assumption.

either meaningful discouragement or reinforcement of subsequent vaccination behavior, in line with Schneider et al. (2023).

Using Swedish registry data, we also examine spillover effects on unincentivized family members. The Swedish Tax Agency identified participants' registered or married partners, parents, and children, and the Public Health Agency matched these individuals to complete COVID-19 vaccination records through December 2023. Partners of participants assigned to the guaranteed, lottery, donation, and choice conditions are 7.63 pp ($p = 0.093$), 8.13 pp ($p = 0.084$), 6.24 pp ($p = 0.188$), and 6.21 pp ($p = 0.181$) more likely to vaccinate within 30 days, respectively (see Appendix Section D.5). Pooling across incentive arms, vaccination uptake is 7.05 pp higher than in the control condition ($p = 0.048$). Although imprecisely estimated, these spillover effects are large, at roughly half the size of the direct effects. Vaccination rates among partners remain higher in the long run—pooling across incentives, rates after 660 days are 3.50 pp higher than in the control condition—but are no longer statistically significant. We find no statistically significant effects for children or parents. A likely explanation is coordination: 30.2% of participants received the vaccine on the same day as their partners, compared to only 2.6% with parents and 1.4% with children, likely reflecting differences in vaccination schedules across age groups.

4 Do surrogates predict long-run treatment effects?

In this section, we study whether long-run effects can be predicted using surrogate outcomes. Our long-run outcome of interest is vaccination within 660 days after the experiment. The surrogates we consider are stated intentions to vaccinate, clicks on a link providing information on how to schedule a vaccination appointment, and shorter-run vaccination, such as vaccination within 15, 30, and 150 days. These outcomes cover the range of surrogates typically used in the literature, including self-reported measures, survey behaviors, and short-run administrative data. We first introduce methods to derive such predictions and then evaluate their performance.

4.1 Framework and surrogate methods

We consider a standard surrogacy setting in which researchers have access to experimental data with exogenous treatment assignment, $W_i \in \{1, 0\}$, and observe surrogate outcomes,

S_i, and pre-treatment covariates, X_i, but do not observe the outcome of interest, Y_i. The goal is to recover the average treatment effect, adjusted for pre-treatment variables, on Y_i:

$$\tau \equiv \mathbb{E}\Big[\mathbb{E}[Y_i \mid W_i = 1, X_i] - \mathbb{E}[Y_i \mid W_i = 0, X_i]\Big]. \tag{1}$$

We evaluate two main approaches used in the literature. First, a large empirical literature treats surrogate outcomes as valid proxies when existing evidence from auxiliary datasets indicates that these potential surrogates are correlated with the outcome of interest. This approach essentially equates treatment effects on the surrogate with treatment effects on the outcome of interest:

$$\tau_n \equiv \mathbb{E}\Big[\mathbb{E}[S_i \mid W_i = 1, X_i] - \mathbb{E}[S_i \mid W_i = 0, X_i]\Big]. \tag{2}$$

Second, an emerging literature clarifies the assumptions required for surrogacy approaches to be valid and proposes econometric methods that use auxiliary data to improve predictions of treatment effects.[25] In particular, Athey et al. (2025) considers a setting in which, in addition to the experimental dataset with surrogate outcomes, there is an observational dataset containing both surrogate outcomes (S_i) and long-run outcomes (Y_i). Let $P_i \in \{E, O\}$ indicate whether observation i belongs to the experimental (E) or observational (O) sample. To infer long-run effects, these methods estimate how surrogate outcomes predict long-run outcomes in the observational data, and then use this relationship to convert treatment effects on the surrogate into predicted effects on long-run outcomes:

$$\begin{aligned}\tau_s \equiv \mathbb{E}\Big[\{&\mathbb{E}[\mathbb{E}[Y_i \mid S_i, X_i, P_i = O] \mid W_i = 1, X_i, P_i = E] \\ &- \mathbb{E}[\mathbb{E}[Y_i \mid S_i, X_i, P_i = O] \mid W_i = 0, X_i, P_i = E]\} \;\Big|\; P_i = E\Big].\end{aligned} \tag{3}$$

[25] Our setting corresponds to the classical surrogate outcomes framework, in which researchers have experimental data with only surrogate measures and aim to learn about an outcome of interest (e.g., Prentice 1989; Fleming and DeMets 1996; Weir and Walley 2006; VanderWeele 2013; Anderer, Bastani and Silberholz 2021; Chen and Ritzwoller 2023; Athey et al. 2025). Athey et al. (2025) provide state-of-the-art methods for settings in which researchers additionally have access to auxiliary observational data containing both surrogate outcomes and outcomes of interest, but no information on treatment assignment. Their approach has been used to predict long-run treatment effects in RCTs (e.g., Dynarski et al. 2021; Dal Bó et al. 2021) and extended to other settings (e.g., Battocchi et al. 2021; Yang et al. 2024*a*). A related literature studies settings with observational data that include non-exogenous treatment assignment and uses auxiliary experimental data to recover causal effects (e.g., Van Goffrier, Maystre and Gilligan-Lee 2023; Yang et al. 2024*b*; Obradović 2024; Athey, Chetty and Imbens 2025; Imbens et al. 2025), which involves a different identification problem.

For this approach to be valid, the "surrogacy assumption"—which requires that all effects of the treatment on the outcome of interest operate through the surrogate—must hold. Formally: $W_i \perp\!\!\!\perp Y_i \mid S_i, X_i$.[26] This means that conditioning on S_i renders treatment W_i independent of Y_i, implying that the treatment does not alter the mapping between S_i and Y_i.[27]

Our dataset allows us to both apply and evaluate these surrogacy methods, as it contains experimental data with surrogates (S_i)—intentions, link clicks, and short-run vaccination—as well as long-run vaccination (Y_i), treatment assignment (W_i), and pre-treatment covariates (X_i). We estimate τ_n using the regression framework described in Section 2.5 and refer to it as the "naïve estimator." We estimate τ_s in three steps:

(i) Estimate the relationship between long-run vaccination and the surrogate using data from the control group: $Y_i = \alpha_0 + \beta S_i + X_i'\kappa + \varepsilon_i$ for $W_i = 0$.[28]

(ii) Use the estimated coefficients to predict long-run vaccination for all observations in the full sample: $\widehat{Y}_i^S = \hat{\alpha}_0 + \hat{\beta} S_i + X_i'\hat{\kappa}$ for all i.

(iii) Regress the surrogate-predicted long-run vaccination outcome on treatment indicators: $\widehat{Y}_i^S = \gamma + \tau_s W_i + X_i'\kappa + u_i$.

We refer to this estimator as the "surrogacy estimator." This procedure corresponds to the surrogacy index estimator of Athey et al. (2025).[29] Crucially, surrogacy-based estimators rely on the surrogacy assumption, which we can also test directly by estimating:

$$Y_i = \gamma + \tau_t W_i + \eta S_i + X_i'\kappa + u_i, \tag{4}$$

[26]The surrogacy assumption originates from Prentice (1989). In addition to surrogacy, identification requires causal identification of treatment effects on the surrogate and comparability between the observational and experimental datasets (referred to as unconfoundedness and comparability in Athey et al. 2025). In our setting, both conditions are satisfied: treatment effects are identified through random assignment, and both surrogate and long-run outcomes are observed within the same sample. This allows us to directly assess the validity of the surrogacy assumption.

[27]The surrogacy assumption is analogous to a mediation framework with full mediation by S_i (see, e.g., Huber 2020), implying no direct effect of the treatment on Y_i.

[28]Hence, we use the control group as the observational data set; results are robust to using data from other treatment conditions instead, as well as to an alternative approach that randomly draws 20% of the full sample as the observational data set to generate predictions for the remaining 80% (see Appendix Table C.3).

[29]Appendix Table C.4 shows that our results are robust to using alternative estimators.

where we test the null hypothesis that $\tau_t = 0$. A rejection implies that treatment assignment W_i has residual predictive power for Y_i even after conditioning on S_i, indicating a violation of the surrogacy assumption.

4.2 Prediction of long-run effects using surrogate measures

We find that all the surrogate measures—vaccination intentions, link clicks, and short-run vaccination—are predictive of actual vaccination uptake (see Appendix Section C.1). For example, among participants in the control condition, those reporting an intention to vaccinate are 34.7 and 42.1 pp more likely to vaccinate within 30 days and 660 days, respectively, than those with no such intention ($p < 0.001$). These strong correlations are often interpreted as evidence that such measures constitute valid surrogates.

Our data allow us to directly assess the validity of these surrogates. We assess surrogate performance using several policy-relevant criteria. At a minimum, a useful surrogate should recover the correct qualitative conclusion about the long-run treatment effects, in particular the sign and relative ranking of treatment effects. A more demanding criterion is whether the surrogate recovers the magnitude of the predicted long-run treatment effects. We place particular emphasis on this latter criterion, as the surrogate outcomes literature is primarily concerned with accurately predicting treatment effect magnitudes, which are often central for policy design. In Section 5, we further examine whether surrogate-based predictions lead to the same cost-effectiveness conclusions as the true long-run outcome.

Table 2 provides results using both the naïve estimator and the surrogacy estimator based on intentions, link clicks, and vaccination rates measured at 15, 30, and 150 days after trial participation. Figure 3 complements this analysis by showing results for a broader set of time windows, considering daily vaccination rates from 1 to 300 days.

Panel i) of Table 2 reports results for the naïve estimator, which simply equates long-run effects with treatment effects on the surrogate. Comparing intentions (column 1) to actual vaccination (column 7) reveals large differences. Despite large and statistically significant effects of the guaranteed and lottery incentives on both short- and long-run vaccination, the corresponding effects on intentions are small and statistically insignificant: 2.5 pp ($p = 0.157$, 95% CI: $[-0.1, 6.0]$) and -0.5 pp ($p = 0.796$, 95% CI: $[-4.1, 3.2]$), respectively. For the donation condition, intentions even indicate a statistically significant negative effect of 4 pp ($p = 0.036$, 95% CI: $[-7.8, -0.25]$). Thus, treatment effects on intentions do not reliably

Table 2: Predicted effects based on surrogates

	Predicted Vacc. 660d						Actual Vacc. 660
	Intent.	Click	Index	15d	30d	150d	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
i) Naïve estimator							
Guaranteed Condition	2.53	3.74***	3.14***	9.94***	13.02***	6.66***	8.58***
	(1.79)	(0.90)	(1.03)	(1.94)	(2.11)	(2.17)	(2.14)
Lottery Condition	-0.48	2.83***	1.17	4.06**	8.45***	4.87**	7.18***
	(1.86)	(0.86)	(1.06)	(1.90)	(2.12)	(2.17)	(2.14)
Donation Condition	-4.01**	1.95**	-1.03	3.21*	4.66**	0.24	0.61
	(1.92)	(0.83)	(1.08)	(1.92)	(2.12)	(2.21)	(2.20)
Choice Condition	3.34*	3.66***	3.50***	6.71***	10.90***	5.94***	6.83***
	(1.82)	(0.88)	(1.05)	(1.95)	(2.13)	(2.19)	(2.16)
ii) Surrogacy estimator							
Guaranteed Condition	1.20	0.21	1.10	5.62***	8.52***	6.19***	8.58***
	(0.82)	(0.41)	(0.89)	(1.06)	(1.30)	(1.96)	(2.14)
Lottery Condition	-0.23	0.16	-0.30	2.30**	5.53***	4.52**	7.18***
	(0.87)	(0.32)	(0.91)	(1.09)	(1.33)	(2.02)	(2.14)
Donation Condition	-1.90**	0.11	-1.96**	1.82*	3.05**	0.22	0.61
	(0.89)	(0.22)	(0.92)	(1.06)	(1.32)	(2.05)	(2.20)
Choice Condition	1.58*	0.20	1.49*	3.79***	7.13***	5.52***	6.83***
	(0.85)	(0.40)	(0.89)	(1.11)	(1.28)	(2.05)	(2.16)
Pre-reg. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,314	5,314	5,314	5,324	5,324	5,324	5,324

Note: Columns (1) to (6) present predictions based on different surrogate measures. We consider intentions to vaccinate (column 1), clicking a link with information on how to schedule a vaccination appointment (column 2), a combination of both (column 3), and vaccination rates after 15, 30, and 150 days (columns 4 to 6). Column (7) reports the actual treatment effects on vaccination rates after 660 days. Panels i) and ii) report results based on the naïve estimator and the surrogacy estimator, respectively, as described in Section 4.1. For the naïve estimator, column (3) uses the average of intention and clicking behavior; for the surrogacy estimator, both surrogates can be used jointly without prior aggregation. Data on intentions and clicking behavior are missing for 10 participants; accordingly, the number of observations is slightly lower in columns (1) to (3). All analyses include the pre-registered set of controls. Heteroskedasticity-robust standard errors (naïve estimator, Actual Vacc. 660) and bootstrap standard errors (surrogacy estimator; 500 repetitions) are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

reflect the magnitude of the effects on vaccination uptake nor the relative ranking of treatment effects, and, in some cases, even fail to recover the correct sign.

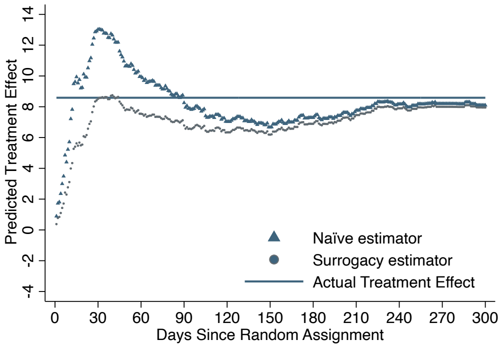
Using clicks on the appointment link as a surrogate performs somewhat better, capturing the sign and qualitative ranking of treatments (column 2). However, the magnitudes remain far from the actual effects, both in the short and long run. Combining clicks and intentions into an index does not resolve these issues: magnitudes remain attenuated and the sign is again incorrect for the donation condition (column 3).

Finally, estimates based on short-run vaccination rates, reported in columns (4) to (6), vary substantially across time windows. At 30 days, they tend to overestimate treatment effects, while at 150 days they underestimate them. More generally, the implied effects fluctuate markedly across horizons and only converge to the long-run effects after roughly 300 days (see Figure 3). While magnitudes vary substantially, these short-run results generally provide a good qualitative assessment of the long-run results.

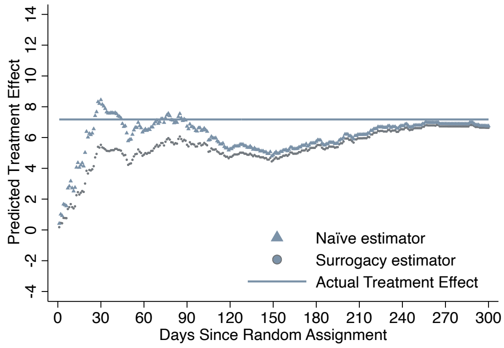
Surrogacy methods aim to improve predictions by scaling treatment effects on surrogates using data on the relationship between the surrogate and the outcome of interest. Panel ii) of Table 2 reports results from the surrogacy estimator. While this approach brings predictions closer to the actual effects, noticeable discrepancies remain. In particular, survey-based surrogates—intentions and link clicks—continue to substantially underestimate treatment effects (columns 1 to 3). For vaccination rates, performance depends on the time window: the estimator performs well for some windows, such as 30 days, but over- or underestimates effects for many shorter and longer horizons (columns 4 to 6). Figure 3 illustrates this pattern across all daily windows between 1 and 300 days. Only at very long horizons do predictions become accurate, as vaccination rates mechanically converge to their long-run levels and there is little remaining variation in the outcome (see Figure 2). This convergence does not reflect predictive power of the surrogate, but rather the fact that the surrogate and the outcome coincide at long horizons. For many shorter horizons, short-run vaccination rates remain unreliable predictors of long-run effects.

A direct test of the surrogacy assumption provides the most stringent validation. This assumption is central to the surrogacy framework: if it fails, predictions based on surrogate outcomes are generally invalid. Our data allow us to test this assumption directly, as described in Section 4.1. Table 3 reports results from regressions of long-run vaccination (after 660 days) on surrogates and treatment indicators. Under the surrogacy assumption, treatment

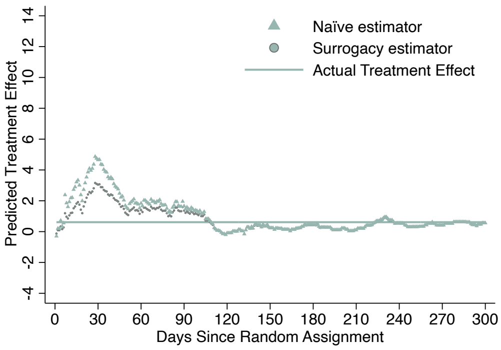
Figure 3: Predicted effects based on vaccination in days since assignment



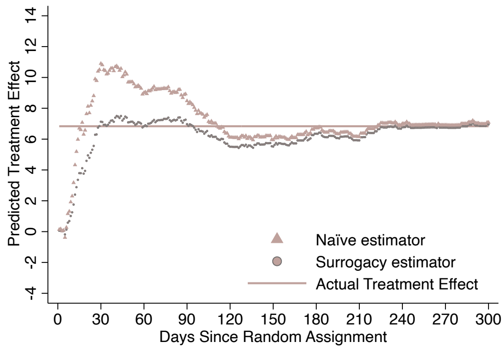
(a) Guaranteed



(b) Lottery



(c) Donation



(d) Choice

Note: This figure presents predictions of the long-run treatment effects (Vaccination after 660 days) for each condition based on vaccination rates from 1 to 300 days. Results are based on the naïve estimator and the surrogacy estimator, as described in Section 4.1.

indicators should have no predictive power once conditioning on the surrogate. We reject this implication for most interventions when using intentions, link clicks, the index using intentions and link clicks, and vaccination measured at 15 and 150 days. This indicates that treatment affects long-run vaccination through channels not captured by these surrogates. Taken together, the results show that surrogate-based predictions can be misleading, and that the assumptions required for their validity are violated in our setting.

So far, we have focused on predicting long-run outcomes. For survey-based surrogates, one might expect better performance when predicting short-run vaccination. However, this is not the case. Using the naïve estimator (Table 2), estimated effects based on intentions and link clicks do not align with short-run vaccination outcomes either. Appendix Tables C.2 and C.5 further show that these measures also perform poorly applying the surrogacy estimator, and that the surrogacy assumption is again rejected. These findings suggest that such surrogates are generally ill-suited for predicting treatment effects on actual vaccination behavior.[30]

Our setting allows us to shed some light on why the surrogacy assumption fails. For link clicks, engagement with the link is very low, ranging from 2.6% in the control group to 6.1% in the choice condition, making it unlikely that treatment effects operate exclusively through this channel. This points to a broader limitation of behavioral proxies that require active engagement: many individuals do not incur even small effort costs, which can then lead to violations of the surrogacy assumption.

This concern is less relevant for self-reported intentions, as all participants are required to respond, and expressing one intention rather than another does not entail differential effort costs. To understand why the surrogacy assumption is violated in this case, Appendix Table C.6 regresses vaccination outcomes on treatment indicators conditional on intentions. The violations are primarily driven by individuals who report intending to vaccinate. This pattern is particularly clear for vaccination within 30 days, although estimates for longer horizons are noisier. A natural interpretation, consistent with insights from behavioral economics, is that incentives increase follow-through among individuals who already intend to vaccinate but would otherwise fail to act, thereby reducing the intention–behavior gap (Sheeran,

[30] Campos-Mercade et al. (2021) report treatment effects of incentives on both intentions and short-run vaccination rates for the first COVID-19 vaccine dose. Both treatment effects were roughly aligned, potentially suggesting that the naïve estimator would not perform poorly in that context. However, the study did not assess the validity of these measures as surrogates, and our results show that intentions are in fact an unreliable surrogate, even for short-run vaccination behavior.

Table 3: Testing the Surrogacy Assumption

	Vacc. 660 conditional on surrogate					
	Intention	Clicked	Intention and Clicked	15 Days	30 Days	150 Days
	(1)	(2)	(3)	(4)	(5)	(6)
Guaranteed Condition	7.44***	7.98***	7.17***	3.03*	-0.19	2.59***
	(1.97)	(2.14)	(1.98)	(1.84)	(1.59)	(0.85)
Lottery Condition	7.37***	6.72***	7.15***	4.91***	1.48	2.80***
	(1.97)	(2.14)	(1.97)	(1.87)	(1.64)	(0.87)
Donation Condition	2.39	0.30	2.23	-1.18	-2.53	0.40
	(2.04)	(2.19)	(2.04)	(1.89)	(1.64)	(0.81)
Choice Condition	5.42***	6.25***	5.16***	3.09*	-0.51	1.49*
	(1.99)	(2.16)	(2.00)	(1.87)	(1.63)	(0.85)
Intention	0.43***		0.43***			
	(0.01)		(0.01)			
Clicked on Appt. Link		0.16***	0.08***			
		(0.03)	(0.03)			
Vacc. 15 Days				0.56***		
				(0.01)		
Vacc. 30 Days					0.67***	
					(0.01)	
Vacc. 150 Days						0.90***
						(0.01)
Pre-reg. controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,314	5,324	5,314	5,324	5,324	5,324

Note: The table shows coefficient estimates from linear regressions of long-run third-dose COVID-19 vaccination uptake on indicators for the experimental conditions, controlling for surrogate measures and the pre-registered set of controls. Under the surrogacy assumption, treatment indicators should have no predictive power once conditioning on the surrogate. Columns (1) to (6) control for intentions to vaccinate (column (1)), clicking a link with information on how to schedule a vaccination appointment (column (2)), a combination of both (column (3)), and vaccination rates after 15, 30, and 150 days (columns (4) to (6)). All regressions include the pre-registered set of controls. Heteroskedasticity-robust standard errors are shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2002; Milkman et al., 2011; Rogers et al., 2015). Such treatment effects on the mapping between surrogate and outcome constitute a direct violation of the surrogacy assumption. More broadly, self-reported measures are particularly prone to such violations, as any mechanism that distorts the mapping between surrogates and actual behavior—such as experimenter demand effects or social desirability bias—will generate the same issue.

Overall, evaluating these policies based on surrogate outcomes would lead to substantially different conclusions than those based on long-run data, as the key assumptions required for valid prediction are violated in our setting.

5 Cost-effectiveness of incentives for vaccination

For policy-makers choosing among interventions to increase vaccination, a central question is not only the effectiveness of interventions, but also the cost per additional vaccination generated. In this section, we study these costs for the marginal individual, focusing on the long-run treatment effects estimated in Section 3. We also examine the potential mistakes that arise when relying on the predicted treatment effects based on the surrogate outcomes introduced in Section 4.

Evaluating the long-run cost-effectiveness of our interventions requires estimates of their long-run treatment effects and their costs. We measure treatment effects as the additional number of people vaccinated per 100 individuals, $E(T_i)$, and calculate costs for the corresponding group of 100 individuals, C_i. We then evaluate cost-effectiveness using the ratio $C_i/E(T_i)$, which captures the cost per additional long-run vaccination. In our setting, the estimated long-run treatment effects are given in Table 1, and the estimated costs—calculated as the number of individuals who got vaccinated within the 30-day window multiplied by the incentive amounts—are $C_{Guaranteed} = 904$, $C_{Lottery} = 812$, $C_{Donation} = 740$, and $C_{Choice} = 864$. Table D.9 reports the implied cost per additional long-run vaccination. Guaranteed incentives yield the lowest cost per additional vaccination at \$105, followed by lottery and choice conditions at \$113 and \$127, respectively. Thus, the greater long-run effectiveness of guaranteed incentives offsets their higher cost. In contrast, donation incentives are substantially less cost-effective, at \$1,213 per additional vaccinated individual. When accounting for spillovers on partners, costs are slightly lower in each condition (Appendix Table D.9). Overall, across most specifications, guaranteed incentives emerge as the most cost-effective policy, with lottery incentives achieving similar cost-effectiveness and remaining potentially attractive in settings with political or administrative constraints on direct payments.[31]

We next assess whether policy evaluations based on surrogate outcomes would lead to accurate conclusions about cost-effectiveness. Appendix Table D.9 reports the implied cost-effectiveness when $E(T_i)$ is replaced by estimates based on intentions, link clicks, or short-run vaccination. Surrogates based on intentions or link clicks often suggest that treatments are ineffective or even counterproductive, leading to very high implied costs. Short-run vaccination performs somewhat better, but still deviates substantially from the true long-run cost-effectiveness. Overall, neither the naïve nor the surrogacy-based approach recovers the correct ranking or magnitude of policies in terms of cost-effectiveness. Consequently, their usefulness for cost–benefit analysis remains limited, and basing policy decisions on such measures could lead to substantial errors.

The reason that surrogate-based predictions fail to recover the correct cost-effectiveness of policies is that they are often unreliable in predicting the magnitude of treatment effects. However, some surrogates, such as link clicks or short-run vaccination, still recover the sign of treatment effects and sometimes approximate the ranking of policies. This raises the question of whether such information can be sufficient for policy design. When costs are identical across policies, ranking treatment effects is sufficient to recover the cost-effectiveness ordering. This may hold in some settings—for example, nudge interventions that differ only in the messages sent to participants (e.g., Milkman et al. 2021*b*). In such cases, a surrogate that correctly captures the sign and ranking of treatment effects, such as link clicks, can be useful for policy design. However, once policies differ in cost, as in our setting, this correspondence breaks down. Indeed, although link clicks correctly rank treatments by effectiveness, they fail to recover the correct ranking in terms of cost-effectiveness. Naturally, rankings based solely on

[31] Existing estimates typically place the benefits per additional vaccination generated on the order of a few hundred US dollars per dose, although there is meaningful variation depending on the population targeted, the epidemiological context, and the timing of vaccination (Bartsch et al., 2021; Barro, 2022; López et al., 2022; Sevilla et al., 2024; Utami et al., 2023). For example, using standard valuations of health improvements, recent work using US data from 2023–2024 estimates direct health benefits in the range of roughly $90 to $360 per additional COVID-19 booster dose (Prosser et al., 2025). These estimates suggest that guaranteed, lottery and choice incentives are likely to be socially beneficial. The evaluation of the donation condition, however, depends on how the generated donations are used; such a policy could also be attractive, but only if the social value of the donations greatly exceeds the benefit of providing equivalent transfers directly to individuals.

treatment effects are insufficient when costs vary across interventions, which highlights the limitation of surrogates only providing the correct ranking.[32]

6 Conclusion

If effective, incentives could be a powerful tool to promote health behaviors that benefit society and reduce the high burden of disease linked to harmful behaviors (WHO, 2023). Yet most existing evidence relies on surrogate outcomes—such as intentions or short-run behavioral responses—rather than the long-run behavior that is often most relevant for policy. Using a large-scale field experiment with linked administrative vaccination records, we study whether incentives have persistent effects on vaccination and whether commonly used surrogates can predict these effects.

We find that financial incentives substantially increase vaccination uptake and generate durable gains. Guaranteed payments of $20 raise vaccination by 13 percentage points in the short run and by more than 8 percentage points in the long run. Lottery and choice incentives also produce lasting effects, while donation-based incentives mainly accelerate uptake. These effects are observed across socio-demographic groups, including individuals with negative vaccination attitudes, and do not crowd out future vaccination behavior.

We believe that these findings are likely to extend beyond our setting. Sweden is comparable to other high-income countries in vaccination rates, inequality, and economic preferences (Falk et al., 2018; WGI, 2023; Mathieu et al., 2020), and COVID-19 booster decisions closely resemble other recurrent vaccination settings, such as flu, measles, or human papillomavirus. Moreover, although the distribution of individuals may differ across countries, the consistent effectiveness of incentives across groups suggests a generalizable pattern in response to incentives. At the same time, our study focuses on a single incentive size and one-time payments, leaving open how alternative designs—such as different incentive sizes or repeated payments—would affect vaccination behavior.

Beyond documenting the long-run effects of incentives, our study also sheds light on the performance of commonly used surrogate outcomes. In much of the empirical literature in the

[32] Correctly identifying the sign of an intervention may still be valuable. Policies that reduce vaccination uptake cannot be cost-effective, so behavioral proxies that reliably recover the direction of effects may help rule out harmful interventions early on (e.g., before scaling or running larger trials). Similarly, if costs do not matter much in an emergency setting, ranking policies by their likely effectiveness would still be useful for policy-makers to quickly identify the most promising intervention.

social sciences and public health, researchers rely on such surrogates—often because long-run behavioral data are unavailable, costly, or slow to obtain—and typically justify their use on the basis that these measures are correlated with the outcome of interest. In our study, short-run measures such as stated intentions, engagement with appointment links, or early vaccination are indeed strongly correlated with eventual vaccination. However, our results show that these correlations do not necessarily translate into predictive validity for treatment effects. In fact, intention data can misrepresent even the direction of long-run impacts. Applying state-of-the-art surrogacy methods confirms that the assumptions required for valid prediction are frequently violated. As a result, neither stated intentions nor engagement with appointment links reliably predict the magnitude of long-run effects, while short-run vaccination performs somewhat better but remains sensitive to the time window.

The breadth of our data further allows us to offer insights into why the assumptions behind surrogate-based inference fail. In particular, we find that participants who report an intention to vaccinate are more likely to follow through when incentives are present. This implies that incentives not only shift surrogate outcomes, but also alter the mapping between surrogates and actual behavior, violating the conditions required for valid prediction. This mechanism is consistent with evidence from behavioral economics showing that interventions can affect the translation of intentions into actions (e.g., Sheeran, 2002; Milkman et al., 2011; Rogers et al., 2015), for instance by reducing frictions or increasing follow-through. Other factors, such as experimenter demand effects (Zizzo, 2010; De Quidt, Haushofer and Roth, 2018) and social desirability (Nederhof, 1985), could similarly violate the surrogacy assumption.

Our paper therefore lies at the intersection of behavioral and experimental economics and the econometrics literature on surrogate outcomes. It contributes to both literatures and shows how they can complement each other. By applying methods developed in the econometrics literature, we show that common validation approaches in experimental economics—based on correlations between surrogate and long-run outcomes—can be misleading. At the same time, insights from behavioral economics help explain when and why the key assumptions underlying surrogacy methods may fail.

Overall, our findings highlight the importance of carefully evaluating the assumptions required to generalize from surrogate outcomes. In particular, they caution against over-reliance on survey-based or short-run proxies when evaluating interventions, especially in contexts where long-run behavior is the ultimate policy-relevant outcome.

References

Alsan, Marcella, and Sarah Eichmeyer. 2024. "Experimental evidence on the effectiveness of nonexperts for improving vaccine demand." *American Economic Journal: Economic Policy*, 16(1): 394–414.

Alsan, M., O. Garrick, and G. Graziani. 2019. "Does diversity matter for health? Experimental evidence from Oakland." *American Economic Review*, 109(12): 4071–4111.

Amabile, T. M., K. G. Hill, B. A. Hennessey, and E. M. Tighe. 1994. "The Work Preference Inventory: assessing intrinsic and extrinsic motivational orientations." *Journal of Personality And Social Psychology*, 66(5): 950.

Anderer, Arielle, Hamsa Bastani, and John Silberholz. 2021. "Adaptive Clinical Trial Designs with Surrogates: When Should We Bother?" *Management Science*, 68(3): 1982–2002.

Angelucci, M., S. Prina, H. Royer, and A. Samek. 2019. "Incentives and unintended consequences: Spillover effects in food choice." *American Economic Journal: Economic Policy*, 11(4): 66–95.

Angerer, Silvia, Daniela Glätzle-Rützler, Philipp Lergetporer, and Thomas Rittmannsberger. 2023. "How does the vaccine approval procedure affect COVID-19 vaccination intentions?" *European Economic Review*, 158: 104504.

Aquino, K., and A. Reed II. 2002. "The self-importance of moral identity." *Journal of Personality And Social Psychology*, 83(6): 1423.

Ashworth, Madison, Linda Thunström, Todd L Cherry, Stephen C Newbold, and David C Finnoff. 2021. "Emphasize personal health benefits to boost COVID-19 vaccination rates." *Proceedings of the National Academy of Sciences*, 118(32): e2108225118.

Athey, Susan, and Stefan Wager. 2021. "Policy Learning With Observational Data." *Econometrica*, 89(1): 133–161.

Athey, Susan, Raj Chetty, and Guido Imbens. 2025. "The experimental selection correction estimator: Using experiments to remove biases in observational estimates." National Bureau of Economic Research.

Athey, Susan, Raj Chetty, Guido W Imbens, and Hyunseung Kang. 2025. "The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely." *Review of Economic Studies*, rdaf087.

Banerjee, A. V., E. Duflo, R. Glennerster, and D. Kothari. 2010. "Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives." *Bmj*, 340.

Barankay, Iwan, Peter P. Reese, Mary E. Putt, Louise B. Russell, George Loewenstein, David Pagnotti, Jiali Yan, Jingsan Zhu, Ryan McGilloway, Troyen Brennan, Darra Finnerty, Karen Hoffer, Sakshum Chadha, and Kevin G. Volpp. 2020. "Effect of Patient Financial Incentives on Statin Adherence and Lipid Control: A Randomized Clinical Trial." *JAMA Network Open*, 3(10): e2019429–e2019429.

Barber, A, and J West. 2022. "Conditional cash lotteries increase COVID-19 vaccination rates." *Journal of Health Economics*, 81: 102578.

Barro, Robert J. 2022. "Vaccination rates and COVID outcomes across US states." *Economics & Human Biology*, 47: 101201.

Bartsch, Sarah M, Patrick T Wedlock, Kelly J O'Shea, Sarah N Cox, Ulrich Strych, Jennifer B Nuzzo, Marie C Ferguson, Maria Elena Bottazzi, Sheryl S Siegmund, Peter J Hotez, et al. 2021. "Lives and costs saved by expanding and expediting coronavirus disease 2019 vaccination." *The Journal of infectious diseases*, 224(6): 938–948.

Batteux, E., F. Mills, L. F. Jones, C. Symons, and D. Weston. 2022. "The effectiveness of interventions for increasing COVID-19 vaccine uptake: A systematic review." *Vaccines*, 10(3): 386.

Battocchi, Keith, Eleanor Dillon, Maggie Hei, Greg Lewis, Miruna Oprescu, and Vasilis Syrgkanis. 2021. "Estimating the long-term effects of novel treatments." *Advances in Neural Information Processing Systems*, 34: 2925–2935.

Becker, G. S., and J. J. Elias. 2007. "Introducing incentives in the market for live and cadaveric organ donations." *Journal of Economic Perspectives*, 21(3): 3–24.

Belot, M., J. James, and P. Nolen. 2016. "Incentives and children's dietary choices: A field experiment in primary schools." *Journal of Health Economics*, 50: 213–229.

Betsch, Cornelia, Robert Böhm, Lars Korn, and Cindy Holtmann. 2017. "On the benefits of explaining herd immunity in vaccine advocacy." *Nature human behaviour*, 1(3): 0056.

Bowles, Samuel, and Sandra Polanía-Reyes. 2012. "Economic Incentives and Social Preferences: Substitutes or Complements?" *Journal of Economic Literature*, 50(2): 368–425.

Brewer, Noel T, Gretchen B Chapman, Alexander J Rothman, Julie Leask, and Allison Kempe. 2017. "Increasing vaccination: putting psychological science into action." *Psychological Science in the Public Interest*, 18(3): 149–207.

Bronchetti, Erin Todd, David B. Huffman, and Ellen Magenheim. 2015. "Attention, intentions, and follow-through in preventive health behavior: Field experimental evidence on flu vaccination." *Journal of Economic Behavior And Organization*, 116: 270–291.

Buser, Thomas, Muriel Niederle, and Hessel Oosterbeek. 2024. "Can competitiveness predict education and labor market outcomes? Evidence from incentivized choice and survey measures." *Review of Economics and Statistics*, 1–45.

Campos-Mercade, Pol, Armando N. Meier, Devin Pope, and Florian H. Schneider. 2023. "Motivating vaccination with financial incentives." *Trends in Cognitive Sciences*, 27(12): 1099–1101.

Campos-Mercade, Pol, Armando N. Meier, Florian H. Schneider, Stephan Meier, Devin Pope, and Erik Wengström. 2021. "Monetary incentives increase COVID-19 vaccinations." *Science*, 374(6569): 879–882.

Carrera, Mariana, Heather Royer, Mark Stehr, Justin Sydnor, and Dmitry Taubinsky. 2022. "Who Chooses Commitment? Evidence and Welfare Implications." *The Review of Economic Studies*, 89(3): 1205–1244.

Carrera, M., H. Royer, M. Stehr, and J. Sydnor. 2020. "The structure of health incentives: Evidence from a field experiment." *Management Science*, 66(5): 1890–1908.

Cavatorta, Elisa, and David Schröder. 2019. "Measuring ambiguity preferences: A new ambiguity preference survey module." *Journal of Risk and Uncertainty*, 58(1): 71–100.

Chang, Tom Y, Mireille Jacobson, Manisha Shah, Rajiv Pramanik, and Samir B Shah. 2026. "Impact of reminder messages, with and without financial incentives, on influenza vaccination: A randomized trial in a California health system." *Vaccine*, 75: 128263.

Chang, T. Y., M. Jacobson, M. Shah, M. Kopetsky, R. Pramanik, and S. B. Shah. 2023. "Reminders, but not monetary incentives, increase COVID-19 booster uptake." *Proceedings of The National Academy of Sciences*, 120(31): e2302725120.

Chapman, Jonathan, Pietro Ortoleva, Erik Snowberg, Leeat Yariv, and Colin Camerer. 2025. "Reassessing qualitative self-assessments and experimental validation." National Bureau of Economic Research.

Charness, Gary, and Uri Gneezy. 2009. "Incentives to Exercise." *Econometrica*, 77(3): 909–931.

Chen, Jiafeng, and David M Ritzwoller. 2023. "Semiparametric estimation of long-term treatment effects." *Journal of Econometrics*, 237(2): 105545.

Cohn, E., M. Chimowitz, T. Long, J. K. Varma, and D. A. Chokshi. 2022. "The effect of a proof-of-vaccination requirement, incentive payments, and employer-based mandates on COVID-19 vaccination rates in New York City: A synthetic-control analysis." *The Lancet Public Health*, 7(9): e754–e762.

Dai, Hengchen, Silvia Saccardo, Maria A Han, Lily Roh, Naveen Raja, Sitaram Vangala, Hardikkumar Modi, Shital Pandya, Michael Sloyan, and Daniel M Croymans. 2021. "Behavioural nudges increase COVID-19 vaccinations." *Nature*, 597(7876): 404–409.

Dal Bó, Ernesto, Frederico Finan, Nicholas Y Li, and Laura Schechter. 2021. "Information technology and government decentralization: Experimental evidence from paraguay." *Econometrica*, 89(2): 677–701.

DellaVigna, S., and D. Pope. 2018. "What motivates effort? Evidence and expert forecasts." *The Review of Economic Studies*, 85(2): 1029–1069.

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and bounding experimenter demand." *American Economic Review*, 108(11): 3266–3302.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner. 2011. "Individual risk attitudes: Measurement, determinants, and behavioral consequences." *Journal of the european economic association*, 9(3): 522–550.

Dolan, P., M. M. Galizzi, and D. Navarro-Martinez. 2015. "Paying people to eat or not to eat? Carryover effects of monetary incentives on eating behaviour." *Social Science & Medicine*, 133: 153–158.

Duch, Raymond M, Adrian Barnett, Maciej Filipek, Javier Espinosa-Brito, Laurence SJ Roope, Mara Violato, and Philip M Clarke. 2023*a*. "Cash versus lottery video messages: online COVID-19 vaccine incentives experiment." *Oxford Open Economics*, 2: odad004.

Duch, R., E. Asiedu, R. Nakamura, T. Rouyard, A. Mayol, A. Barnett, and P. Clarke. 2023*b*. "Financial incentives for COVID-19 vaccines in a rural low-resource setting: a cluster-randomized trial." *Nature Medicine*, 29(12): 3193–3202.

Dynarski, Susan, Carmello J Libassi, Katherine Michelmore, and Stephanie Owen. 2021. "Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students." *American Economic Review*, 111(6): 1721–1756.

Enke, Benjamin, Ricardo Rodriguez-Padilla, and Florian Zimmermann. 2022. "Moral universalism: Measurement and economic relevance." *Management Science*, 68(5): 3590–3603.

Fabbri, M., P. Nicola Barbieri, and M. Bigoni. 2019. "Ride your luck! A field experiment on lottery-based incentives for compliance." *Management Science*, 65(9): 4336–4348.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global Evidence on Economic Preferences." *Quarterly Journal of Economics*, 1645–1692.

Falk, Armin, Anke Becker, Thomas Dohmen, David Huffman, and Uwe Sunde. 2023. "The preference survey module: A validated instrument for measuring risk, time, and social preferences." *Management Science*, 69(4): 1935–1950.

Fleming, Thomas R, and David L DeMets. 1996. "Surrogate end points in clinical trials: are we being misled?" *Annals of internal medicine*, 125(7): 605–613.

Fuster, Andreas, and Basit Zafar. 2023. "Survey experiments on economic expectations." In *Handbook of Economic Expectations*. 107–130. Elsevier.

Giuntella, O., S. Saccardo, and S. Sadoff. 2024. “Sleep: Educational Impact and Habit Formation.”

Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. “When and Why Incentives Don’t Work to Modify Behavior.” *Journal of Economic Perspectives*, 25(4): 191–210.

Goette, Lorenz, and Alois Stutzer. 2020. “Blood donations and incentives: Evidence from a field experiment.” *Journal of Economic Behavior & Organization*, 170: 52–74.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2023. “Designing information provision experiments.” *Journal of economic literature*, 61(1): 3–40.

Huang, Y., X. Huang, and R. Yu. 2023. “The effectiveness of nonfinancial interventions and monetary incentives on COVID-19 vaccination: A meta-analysis.” *Health Psychology*, 42(6): 411.

Huber, Martin. 2020. “Mediation analysis.” *Handbook of labor, human resources and population economics*, 1–38.

Imas, A. 2014. “Working for the “warm glow”: On the benefits and limits of prosocial incentives.” *Journal of Public Economics*, 114: 14–18.

Imbens, Guido, Nathan Kallus, Xiaojie Mao, and Yuhao Wang. 2025. “Long-term causal inference under persistent confounding via data combination.” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 87(2): 362–388.

Jacobson, M., T. Y. Chang, M. Shah, R. Pramanik, and S. B. Shah. 2022. “Can financial incentives and other nudges increase COVID-19 vaccinations among the vaccine hesitant? A randomized trial.” *Vaccine*, 40(43): 6235–6242.

Jarrett, Caitlin, Rose Wilson, Maureen O’Leary, Elisabeth Eckersberger, Heidi J Larson, et al. 2015. “Strategies for addressing vaccine hesitancy–A systematic review.” *Vaccine*, 33(34): 4180–4190.

Jecker, N. S. 2021. “Cash incentives, ethics, and COVID-19 vaccination.” *Science*, 374(6569): 819–820.

Jones, Damon, David Molitor, and Julian Reif. 2024. “Incentives and Habit Formation in Health Screenings: Evidence from the Illinois Workplace Wellness Study.” *Working Paper.*

Karing, Anne. 2024. “Social signaling and childhood immunization: A field experiment in Sierra Leone.” *The Quarterly Journal of Economics*, 139(4): 2083–2133.

Khazanov, G. K., R. Stewart, M. F. Pieri, C. Huang, C. T. Robertson, K. A. Schaefer, and J. Fishman. 2023. “The effectiveness of financial incentives for COVID-19 vaccination: A systematic review.” *Preventive Medicine*, 107538.

Klüver, Heike, Felix Hartmann, Macartan Humphreys, Ferdinand Geissler, and Johannes Giesecke. 2021. “Incentives can spur COVID-19 vaccination uptake.” *Proceedings of the National Academy of Sciences*, 118(36): e2109543118.

Knaus, M. C. 2022. “Double machine learning-based programme evaluation under unconfoundedness.” *The Econometrics Journal*, 25(3): 602–627.

Lacetera, N., and M. Macis. 2010. “Do all material incentives for pro-social activities backfire? The response to cash and non-cash incentives for blood donations.” *Journal of Economic Psychology*, 31(4): 738–748.

Lacetera, N., M. Macis, and R. Slonim. 2013. “Economic rewards to motivate blood donations.” *Science*, 340(6135): 927–928.

Lacetera, N., M. Macis, and R. Slonim. 2014. “Rewarding volunteers: A field experiment.” *Management Science*, 60(5): 1107–1129.

Largent, Emily A., and Franklin G. Miller. 2021. "Booster Doses for COVID-19 Vaccines: Ethical Considerations." *Jama*, 325(6): 534–535.

Lau, Krystal, Marisa Miraldo, Matteo M Galizzi, and Katharina Hauck. 2019. "Social norms and free-riding in influenza vaccine decisions in the UK: an online experiment." *The Lancet*, 394: S65.

Lieberman, A., A. Gneezy, E. Berry, S. Miller, M. Koch, C. Ahn, and S. Gupta. 2019. "Financial incentives to promote colorectal cancer screening: a longitudinal randomized control trial." *Cancer Epidemiology, Biomarkers & Prevention*, 28(11): 1902–1908.

List, J. A., and A. S. Samek. 2015. "The behavioralist as nutritionist: Leveraging behavioral economics to improve child food choice and consumption." *Journal of Health Economics*, 39: 135–146.

López, Francesc, Martí Català, Clara Prats, Oriol Estrada, Irene Oliva, Núria Prat, Mar Isnard, Roser Vallès, Marc Vilar, Bonaventura Clotet, et al. 2022. "A cost–benefit analysis of COVID-19 vaccination in Catalonia." *Vaccines*, 10(1): 59.

Malik, Amyn A, Noureen Ahmed, Mehr Shafiq, Jad A Elharake, Erin James, Kate Nyhan, Elliott Paintsil, Hannah Camille Melchinger, Yale Behavioral Interventions Team, Fauzia A Malik, et al. 2023. "Behavioral interventions for vaccination uptake: A systematic review and meta-analysis." *Health Policy*, 137: 104894.

Mantzari, E., F. Vogt, and T. M. Marteau. 2015. "Financial incentives for increasing uptake of HPV vaccinations: a randomized controlled trial." *Health Psychology*, 34(2): 160.

Mathieu, Edouard, Hannah Ritchie, Lucas Rodés-Guirao, Cameron Appel, Charlie Giattino, Joe Hasell, Bobbie Macdonald, Saloni Dattani, Diana Beltekian, Esteban Ortiz-Ospina, and Max Roser. 2020. "Coronavirus Pandemic (COVID-19)."

Mellström, Carl, and Magnus Johannesson. 2008. "Crowding Out in Blood Donation: Was Titmuss Right?" *Journal of The European Economic Association*, 6(4): 845–863.

Milkman, Katherine L., Dena Gromet, Hung Ho, Joseph S. Kay, Timothy W. Lee, Pepi Pandiloski, Yeji Park, Aneesh Rai, Max Bazerman, John Beshears, Lauri Bonacorsi, Colin Camerer, Edward Chang, Gretchen Chapman, Robert Cialdini, Hengchen Dai, Lauren Eskreis-Winkler, Ayelet Fishbach, James J. Gross, Samantha Horn, Alexa Hubbard, Steven J. Jones, Dean Karlan, Tim Kautz, Erika Kirgios, Joowon Klusowski, Ariella Kristal, Rahul Ladhania, George Loewenstein, Jens Ludwig, Barbara Mellers, Sendhil Mullainathan, Silvia Saccardo, Jann Spiess, Gaurav Suri, Joachim H. Talloen, Jamie Taxer, Yaacov Trope, Lyle Ungar, Kevin G. Volpp, Ashley Whillans, Jonathan Zinman, and Angela L. Duckworth. 2021*a*. "Megastudies improve the impact of applied behavioural science." *Nature*, 600(7889): 478–483.

Milkman, Katherine L., John Beshears, James J. Choi, David Laibson, and Brigitte C. Madrian. 2011. "Using implementation intentions prompts to enhance influenza vaccination rates." *Proceedings of The National Academy of Sciences of The United States of America*, 108(26): 10415–10420.

Milkman, Katherine L, Linnea Gandhi, Mitesh S Patel, Heather N Graci, Dena M Gromet, Hung Ho, Joseph S Kay, Timothy W Lee, Jake Rothschild, Jonathan E Bogard, et al. 2022*a*. "A 680,000-person megastudy of nudges to encourage vaccination in pharmacies." *Proceedings of the National Academy of Sciences*, 119(6): e2115126119.

Milkman, Katherine L, Mitesh S Patel, Linnea Gandhi, Heather N Graci, Dena M Gromet, Hung Ho, Joseph S Kay, Timothy W Lee, Modupe Akinola, John Beshears, et al. 2021*b*. "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(20): e2101165118.

Milkman, K. L., L. Gandhi, S. F. Ellis, H. N. Graci, D. M. Gromet, R. S. Mobarak, and K. G. Volpp. 2022*b*. "A citywide experiment testing the impact of geographically targeted, high-pay-off vaccine lotteries." *Nature Human Behaviour*, 6(11): 1515–1524.

Moehring, Alex, Avinash Collis, Kiran Garimella, M Amin Rahimian, Sinan Aral, and Dean Eckles. 2023. "Providing normative information increases intentions to accept a COVID-19 vaccine." *Nature Communications*, 14(1): 126.

Nederhof, Anton J. 1985. "Methods of coping with social desirability bias: A review." *European journal of social psychology*, 15(3): 263–280.

Obradović, Filip. 2024. "Identification of long-term treatment effects via temporal links, observational, and experimental data." *arXiv preprint arXiv:2411.04380.*

OWID. 2022. "Julia Murphy and Max Roser." *Our World in Data.* https://ourworldindata.org/internet.

Oza, A. 2021. *Cash for shots? Studies suggest payouts improve vaccination rates.* Science.

Persad, G., E. J. Emanuel, S. Sangenito, A. Glickman, S. Phillips, and E. A. Largent. 2021. "Public perspectives on COVID-19 vaccine prioritization." *Jama Network Open*, 4(4): e217943–e217943.

Plant, E. A., and P. G. Devine. 1998. "Internal and external motivation to respond without prejudice." *Journal of Personality And Social Psychology*, 75(3): 811.

Prentice, Ross L. 1989. "Surrogate endpoints in clinical trials: definition and operational criteria." *Statistics in medicine*, 8(4): 431–440.

Prosser, Lisa A, Megan Wallace, Angela M Rose, Kerra Mercon, Cara B Janusz, Acham Gebremariam, David W Hutton, Andrew J Leidner, Fangjun Zhou, Ismael R Ortega-Sanchez, et al. 2025. "Cost-effectiveness of 2023-2024 COVID-19 vaccination in US adults." *JAMA network open*, 8(8): e2523688.

Rogers, Todd, Katherine L Milkman, Leslie K John, and Michael I Norton. 2015. "Beyond good intentions: Prompting people to make plans improves follow-through on important tasks." *Behavioral Science & Policy*, 1(2): 33–41.

Royer, Heather, Mark Stehr, and Justin Sydnor. 2015. "Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a Fortune-500 company." *American Economic Journal: Applied Economics*, 7(3): 51–84.

Ruggeri, K., F. Stock, S. A. Haslam, V. Capraro, P. Boggio, N. Ellemers, and R. Willer. 2024. "A synthesis of evidence for policy from behavioural science during COVID-19." *Nature*, 625(7993): 134–147.

Saccardo, Silvia, Hengchen Dai, Maria A Han, Sitaram Vangala, Juyea Hoo, and Jeffrey Fujimoto. 2024. "Field testing the transferability of behavioural science knowledge on promoting vaccinations." *Nature Human Behaviour*, 8(5): 878–890.

Sandel, M. J. 2012. *What money can't buy: the moral limits of markets.* Macmillan.

Savulescu, J. 2021. "Good reasons to vaccinate: mandatory or payment for risk?" *Journal of Medical Ethics*, 47(2): 78–85.

Schneider, F. H., P. Campos-Mercade, S. Meier, D. Pope, E. Wengström, and A. N. Meier. 2023. "Financial incentives for vaccination do not have negative unintended consequences." *Nature*, 613(7944): 526–533.

Serra-Garcia, M., and N. Szech. 2023. "Incentives and defaults can increase COVID-19 vaccine intentions and test demand." *Management Science*, 69(2): 1037–1049.

Sevilla, JP, Daria Burnes, Joseph S Knee, Manuela Di Fusco, Moe H Kyaw, Jingyan Yang, Jennifer L Nguyen, and David E Bloom. 2024. "The global health and economic value of COVID-19 vaccination." *BMJ Global Health*, 9(9).

Sheeran, Paschal. 2002. "Intention—behavior relations: a conceptual and empirical review." *European review of social psychology*, 12(1): 1–36.

Shen, Yang, Jingyu Wang, Jian Wang, Stephen Nicholas, Elizabeth Maitland, Min Lv, Tao Yin, and Dawei Zhu. 2024. "Effectiveness of financial incentives on influenza vaccination among older adults in China: A randomized clinical trial." *Clinical Microbiology and Infection*, 30(7): 911–916.

Stantcheva, Stefanie. 2023. "How to run surveys: A guide to creating your own identifying variation and revealing the invisible." *Annual Review of Economics*, 15(1): 205–234.

Steinert, Janina I, Henrike Sternberg, Hannah Prince, Barbara Fasolo, Matteo M Galizzi, Tim Büthe, and Giuseppe A Veltri. 2022. "COVID-19 vaccine hesitancy in eight European countries: Prevalence, determinants, and heterogeneity." *Science advances*, 8(17): eabm9825.

Thirumurthy, H., K. L. Milkman, K. G. Volpp, A. M. Buttenheim, and D. G. Pope. 2022. "Association between statewide financial incentive programs and COVID-19 vaccination rates." *Plos One*, 17(3): e0263425.

Titmuss, R. M. 1970. *The gift relationship. From human blood to social policy.* Macmillan.

Tonin, M., and M. Vlassopoulos. 2015. "Corporate philanthropy and productivity: Evidence from an online real effort experiment." *Management Science*, 61(8): 1795–1811.

Tuckman, B. W. 1991. "The development and concurrent validity of the procrastination scale." *Educational And Psychological Measurement*, 51(2): 473–480.

Utami, Auliasari Meita, Farida Rendrayani, Qisty Aulia Khoiry, Dita Noviyanti, Auliya A Suwantika, Maarten J Postma, and Neily Zakiyah. 2023. "Economic evaluation of COVID-19 vaccination: A systematic review." *Journal of global health*, 13: 06001.

VanderWeele, Tyler J. 2013. "Surrogate measures and consistent surrogates." *Biometrics*, 69(3): 561–565.

Van Goffrier, Graham, Lucas Maystre, and Ciarán Mark Gilligan-Lee. 2023. "Estimating long-term causal effects from short-term experiments and long-term observational data with unobserved confounding." 791–813, PMLR.

Volpp, Kevin G, Andrea B Troxel, Mark V Pauly, Henry A Glick, Andrea Puig, David A Asch, Robert Galvin, Jingsan Zhu, Fei Wan, Jill DeGuzman, Elizabeth Corbett, Janet Weiner, Janet Audrain-McGovern, and The Transdisciplinary. 2009. "A Randomized, Controlled Trial of Financial Incentives for Smoking Cessation." *New England Journal of Medicine*, 360(7): 699–709.

Volpp, Kevin G., Andrea Gurmankin Levy, David A. Asch, Jesse A. Berlin, John J. Murphy, Angela Gomez, Harold Sox, Jingsan Zhu, and Caryn Lerman. 2006. "A Randomized Controlled Trial of Financial Incentives for Smoking Cessation." *Cancer Epidemiology, Biomarkers & Prevention*, 15(1): 12–18.

Volpp, K. G., G. Loewenstein, and A. M. Buttenheim. 2021. "Behaviorally informed strategies for a national COVID-19 vaccine promotion program." *Jama*, 325(2): 125–126.

Volpp, K. G., L. K. John, A. B. Troxel, L. Norton, J. Fassbender, and G. Loewenstein. 2008. "Financial incentive–based approaches for weight loss: a randomized trial." *Jama*, 300(22): 2631–2637.

Wager, S, and S Athey. 2018. "Estimation and inference of heterogeneous treatment effects using random forests." *Journal of The American Statistical Association*, 113(523): 1228–1242.

Weir, Christopher J, and Rosalind J Walley. 2006. "Statistical evaluation of biomarkers as surrogate endpoints: a literature review." *Statistics in medicine*, 25(2): 183–203.

WGI. 2023. "CAF world giving index 2023."

WHO. 2023. "The contribution of behavioural science to addressing the social and wider determinants of health: evidence review." *Michael P. Kelly, Ananya Arora, Anjuli Banerjee, Jack M. Birch.*

Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the workplace, investment in human capital, and gender." *The Quarterly Journal of Economics*, 133(1): 457–507.

Yang, Jeremy, Dean Eckles, Paramveer Dhillon, and Sinan Aral. 2024*a*. "Targeting for long-term outcomes." *Management Science*, 70(6): 3841–3855.

Yang, Zeqin, Weilin Chen, Ruichu Cai, Yuguang Yan, Zhifeng Hao, Zhipeng Yu, Zhichao Zou, Jixing Xu, Zhen Peng, and Jiecheng Guo. 2024*b*. "Estimating long-term heterogeneous dose-response curve: Generalization bound leveraging optimal transport weights." *arXiv preprint arXiv:2406.19195.*

Zhou, Z., S. Athey, and S. Wager. 2023. "Offline multi-action policy learning: Generalization and optimization." *Operations Research*, 71(1): 148–183.

Zizzo, Daniel John. 2010. "Experimenter demand effects in economic experiments." *Experimental Economics*, 13(1): 75–98.