When nudges aren't enough: Incentives and habit formation in public transport usage

Christina Gravert* and Linus Olsson Collentine

June 25, 2020

Abstract

In three large-scale field experiments with over 32,500 individuals, we investigate whether public transport uptake can be influenced by behavioral interventions and by economic incentives. Despite their effectiveness in other domains, we find a tightly estimated zero for social norms and implementation intentions on ridership. Increasing the economic incentive significantly increases uptake and long-term usage. This increase is sustained for months after removing the incentive. The effect is mainly driven by initial low users, which is evidence for habit formation and highlights the heterogeneous effects of the policy. While there is scope for long-term behavior change, nudging might not be the right approach.

Keywords: transport, nudging, field experiment, habit formation
JEL codes: C93, D04, D91, L91

*Corresponding Author: Christina Gravert, University of Copenhagen, Department of Economics, Øster Farimagsgade 5, 1353 Copenhagen, Denmark; Tel: +45 35 33 47 82, Email: cag@econ.ku.dk. Linus Olsson Collentine, Nudgingbolaget, Email: linus@nudgingbolaget.se. Declarations of interest: none. This paper greatly benefited from comments by John List, Magne Mogstad, Miriam Wüst, participants at the Advances in Field Experiments 2018, Behavioral Public Policy Workshop 2019 in Copenhagen and various seminar participants. We thank Yvonne Werkelid at Skånetrafiken for the collaboration and Isabel Skak Olufsen for research assistance. Gravert was partially funded by the Swedish Environmental Protection Agency. We gratefully acknowledge support from the Center for Economic Behavior and Inequality (CEBI) at the University of Copenhagen, financed by grant #DNRF134 from the Danish National Research Foundation.

1 Introduction

Urban transport accounts for around 40% of end-use energy consumption and trends are increasing (IPCC 2014), posing a threat to our climate and health. Maintaining and increasing the public transport market share is on the national agenda of most industrialized nations, especially the European Union (European Commission 2019). While investments in sustainable infrastructure make up a major share of the development plans, these investments need to be complemented with changes in individual travel behavior.

Keeping infrastructure fixed, individual travel behavior can be influenced through two channels - prices and behavioral interventions. Behavioral interventions, also known as nudges (Thaler & Sunstein 2009), are seen as low cost, easily implementable, and less obtrusive compared to conventional steering instruments such as taxes and regulations (Benartzi et al. 2017, Carlsson et al. forthcoming). If they are effective at changing travel behavior while keeping prices constant, they would be a valuable addition to the policy toolkit.

Several overview articles and policy reports suggest making use of behavioral interventions in the public transport sector (Alta 2018, Metcalfe & Dolan 2012, Garcia-Sierra et al. 2015). While the arguments are persuasive that nudges can correct for inattention, make social norms more salient or help with planning, there is close to no evidence of behavioral interventions affecting public transport usage. In a meta-analysis of 107 Transportation Demand Management studies Ortmann & Dixit (2017) found no sufficiently controlled study on nudges to affect transportation demand. Most studies focus on economic incentives or economic incentives in connection with behavioral elements that were not independently investigated.

This paper aims to fill this gap by testing whether behavioral interventions, social norms, framing, and implementation intentions, have a positive, long-term effect on public transport usage compared to traditional price instruments.

Together with a local transport agency in Southern Sweden, we conducted three natural field experiments with over 32,500 users. The experiments test three different behavioral interventions and are implemented with different cohorts of a public transport marketing campaign. If a person moves to or within the region of Skåne, Sweden, they, shortly after registering their new address, receive an offer to try the regional public transport for two weeks for free. If they do not respond, they receive one follow up reminder. If they do respond, they are sent a pre-paid, reusable travel card that is valid for two weeks and which can be extended with a monthly ticket or a pre-paid amount after the two weeks are up. We implement our experiments within these three stages of the campaign, but on different

cohorts. This allows us to measure the effect of behavioral interventions on the full, random sample and on sub-samples of individuals that have either selected into or not responded to the first stage of the campaign, while also avoiding multiple treatments of our participants.

In our main experiment, we compare the baseline of receiving an offer for a two-week free travel card to 1) an offer with a nudge in the form of a social norm message honestly stating that "72% of their neighbors are traveling with public transport" in addition to the two-week free travel card and 2) an offer of a four-week free travel card, thus doubling the economic incentive. We then tracked the initial response to the offer, the activation of the card, and usage of the card over the following eight months. We supplement the behavioral data with a short follow up survey. We directly compare a nudge to a traditional economic incentive, thus testing the two major channels prices and behavioral motivations against each other. This comparison is important for two reasons. First, it allows us to understand whether a nudge is the most cost-effective intervention. Even if the price instrument would have the larger effect in behavior, the effect might be too small given the costs. A nudge, which in our setting would be costless to implement compared to the price instrument could have a more favorable cost-benefit ratio. Second, in case the nudge is not successful, while the price instrument is, we can infer that individuals are generally able and willing to change their behavior and that the nudge was not strong enough to stimulate a change in behavior.

The group in our main experiment is drawn from the full population of individuals who moved within a given time frame (N=14282). Potentially, a nudge might be more effective on self-selected sub-samples of the population. We, therefore, carry out two further experiments. In the second experiment (N = 11373), we test the social norm nudge on a group of individuals who did not respond to the initial offer of receiving a free two-week travel card. This sub-sample had received the same offer as the baseline group in experiment one, but they are drawn from a different cohort to avoid treatment contamination. We test a neutral baseline against 1) a positively and 2) a negatively framed social norm. In the final experiment (N = 6969), we test an additional nudge, implementation intentions, on a group of individuals who have responded positively to the initial offer of a two-week free travel card. Again, this is a separate cohort. Statistically, only around 70% of the individuals who respond to the offer also use the card at least once. Given their initial positive response, we assume that these individuals are motivated to use public transport, but might encounter barriers of following through with their intentions. A nudge could thus be beneficial to encourage them to make use of the card. In the letter that accompanies the travel card, we compare a standard text describing how to use the card with a text with a prompt to think

about when to use the card for the first time. In a second treatment, we even add blanks that can be filled in with time and place of first usage. There is no change in economic incentives in the two experiments.

We chose social proof, framing and implementation intentions because of their prominence in the nudging literature Carlsson et al. (forthcoming) and their documented positive effects in the environmental or health domain (Cialdini 2003, Allcott 2011, Milkman et al. 2011). Each intervention is aimed at a different bias: social norms, loss aversion and present bias, respectively.

Our randomized and large-scale design overcomes some of the issues of the few previous studies in this area. Participants in our experiment are not aware that they are part of an experiment, and we use administrative data instead of self-reported data, which reduces potential biases or imperfect recall associated with self-reports.

In all three experiments, we find no evidence of the nudges significantly increasing public transport usage. Neither in the main experiment, nor in the two experiments on the self-selected sub-groups do we find any effect on response to the offer, activation or long-term usage. This result could be due to two reasons. Either, individuals were not interested in changing their travel habits, or the nudges were not effective. The evidence points to the latter. In our main experiment, we find that the economic incentive, doubling the free travel time, increased responses to the offer by 16%, from a baseline of 42.8%. The interest in the four-week offer was thus significantly higher than in the two-week offer. Overall, close to every second recipient of the offer of any card decided to respond. Additional evidence that economic incentives (providing individuals with a free card) could be effective in changing behavior. Since we do not have a control group that received no offer, we cannot compare the card to no economic incentives.

Our next result focuses on the activation of the card. It could be that while the four-week offer increases initial interest, this interest does not carry through to actual usage. This is not the case. Unconditionally on the initial response, we find a significantly higher amount of individuals activating a card in the four-week group.

We follow the group from the main experiment for eight months. Eight months later, the cards from the four-week treatment group are significantly more likely (ca. 33%) to still be in our sample of active cards. We measure the effect after the summer holidays in which individuals' travel habits might change due to vacation and a special summer offer on public transport. While the level effect of the active cards is low - only 11% of the individuals who opted into receiving the card are still using it eight months later - the treatment effect is

significant. These results are in line with previous findings of the habit formation literature (Gneezy et al. 2011). Incentives that last at least for a month are usually more effective in changing long-term behavior than shorter incentives.

However, one might be concerned that the effect does not come from habit formation but from initial differences between individuals who self-select into a four-week travel card and those who self select into a two-week travel card. Our evidence is in line with actual habit formation rather than self-selection of individuals who already use public transport regularly. In a supplementary analysis, we classify individuals into low and high users based on their initial travel behavior in week one. Individuals, who from day one of the trial period, use the card at least two times per day are classified as high users and most likely seem to have an established habit. Individuals who use the card less than two times per day during the first week are classified as low users. These are the individuals we are interested in, as they have the potential to increase their usage. The low users in the two-week baseline are significantly more likely to use the card after the experimental period is over than low users in the four-week treatment. We interpret this as clear evidence for habit formation. The increased usage after the experimental period also increased the revenue of the transport agency, as individuals in the four weeks group were more likely to buy monthly travel cards.

This paper contributes to our understanding of the effectiveness of behavioral interventions compared to price instruments in changing behavior. Environmental policy has seen a surge in studies on behavioral interventions that have shown to reduce household energy consumption (Allcott 2011), household water consumption (Ferraro et al. 2011, Tiefenbeck et al. 2016), meat consumption (Gravert & Kurz 2019, Kurz 2018), and food waste (Kallbekken & Sælen 2013). While a large literature has found significant effects of nudges on behavior in the environmental domain (see Carlsson et al. (forthcoming) for a summary), there is reason to believe that switching between transport options comes at a higher effort and even monetary cost than switching between a meat and vegetarian dish at lunch or reducing shower time by a couple of minutes. Findings from one domain cannot easily be transferred to another. There is also a lack of studies comparing price instruments to nudges for the same outcome measures. An exception is Ito (2015) who find that economic incentives have a longer-term effect on reducing energy consumption than a moral appeal. Our findings are in line with a recent paper by (Kristal & Whillans 2019), who also tested several nudges on transportation usage. Like us, they find no effect of nudges such as letters or a customized travel plan on carpooling or bus travel. Their context is different in several dimensions. First, they are nudging employees of one airport, who all travel to the same

workplace, while we are sampling from 1.5 million individuals of the Swedish population, who live and work (or don't work) in various places across the region. Second, they mostly focus on carpooling while we focus on public transport. Third, they only have short-term outcomes, while we can observe travel behavior for eight months. Since we make use of an existing campaign, we keep the experimental setup as natural as possible, which increases external validity.[1]

Further, our study contributes to the formation and persistence of habits. The economic literature provides several examples of habit formation after providing economic incentives in the health domain (Finkelstein et al. 2007, Charness & Gneezy 2009, Volpp, Pauly, Loewenstein & Bangsberg 2009, Volpp, Troxel, Pauly, Glick, Puig, Asch, Galvin, Zhu, Wan, DeGuzman et al. 2009, John et al. 2011, Cawley & Price 2013, Just & Price 2013, Acland & Levy 2015, Royer et al. 2015) and the savings domain (see for example Schaner (2018)). However, in many cases, the effects are short-lived and usually reversed by vacation or holidays unless they come with additional interventions (Gneezy et al. 2011, John et al. 2011, Royer et al. 2015). In our study, we find a significant persistence of habits even after a two months summer break.

What do our results imply for policymakers trying to increase public transport usage through behavioral interventions or price instruments? We tested several popular and feasible behavioral interventions in a natural and easy replicable setting on a meaningful subset of the population. The interventions we chose to test have been shown to be effective in other environmental policy settings. We find no evidence that the nudges had any effect on either the full sample or on selected sub-samples that were either more or less motivated to change their travel habits. The price instrument leads to a meaningful increase in response rates to the offer of a free-travel card and a significant effect on usage of public transport up to eight months later. Taken together with the fact that the average initial response to the offer was 44.6% and that at the end of the free trial period, there was a sharp drop in usage, individuals are clearly responsive to price, which might not come as a surprise to economists. However, given that 72% of the population in the area already at least occasionally uses public transport, it is promising to see that this share could be increased further with subsidies and that while infrastructure investments are certainly an important piece of the puzzle, there is still room for behavior change through policy. Pure nudges, however, do not seem potent enough to affect travel behavior.

[1] Very similar campaigns to ours have been run in both Gothenburg and Stockholm, Sweden, but not scientifically evaluated.

2 Experimental Design

2.1 The setting and data

The experiments were designed and conducted in cooperation with Skånetrafiken, the administration for public transportation in the county Skåne in southernmost Sweden. Skåne has approximately 1.4 million inhabitants and contains urban areas such as Malmö (with 300,000 inhabitants) as well as rural villages. The public transportation managed by Skånetrafiken is used by 250,000 people daily. It consists of buses, trams and regional trains. To travel by public transport, individuals need an electronic JoJo card (a plastic card that looks like a credit card), which can be either uploaded with a pre-paid amount so that for each trip, the payment is deducted, or it can be uploaded with a monthly card for an unlimited amount of trips per month. Every user needs one card which can be re-used until lost or stolen. A single trip could cost between 35 Swedish Kroner (SEK) within a city to 122 SEK (USD 3.68 / USD 12.84) between cities 1,5 hours apart. Travelers need to scan the card every time they use a bus, tram, or train by holding it against a card reader.[2] Even when they have a monthly pass, they need to register their trip. This feature allows us to track trips both during and after the trial period for both the monthly card users and pre-paid users.

The interventions were designed and sent out as part of a flyer campaign by Skånetrafiken, targeted at people who moved to or within the region. During the time frame of the campaign, shortly after moving to a new address, all individuals in the area receive a flyer from Skånetrafiken. The timing is chosen deliberately to give the interventions the highest chances of success. An ideal time to get people to change their travel habits is when they switch residences, as that "unfreezes" the usual habit of traveling the same route to work and other activities (Verplanken et al. 2008). As Goodwin (1977) has argued "the traveler does not carefully and deliberately calculate anew each morning whether to go to work by car or bus. Such deliberation is likely to occur only occasionally, probably in response to some large change in the situation.". Empirical evidence of the effect of a shock on travel routes comes from Larcom et al. (2017). The authors exploit a strike on the London underground system. They show that for those individuals who need to re-optimize their travel routes due to the strike, a significant number sticks with the new route after the strike period. Given this evidence, we make use of the fact that moving to a new address is a natural break in individuals' travel habits.

[2] This is similar to the metro system in NY, London or Paris. Public transport staff will occasionally patrol the trains and buses to check whether riders have an active ticket. Fines for riding without a validated ticket are ca. USD 100.

The basic flyer, which we will use as our baseline, contained the name of the resident, their closest public transport stop, information about Skånetrafiken, and the offer of a free two-week trial (see Figure A.1 in the Appendix for a picture of the flyer). The market value of the free travel card was 625 Swedish Kronor (ca. USD 65). The trial card covered unlimited trips within the region. The card, which is identical to the ones that could be bought, could be used for the duration of the free travel time and could then be uploaded with either a monthly pass, that had to be renewed every month, or with a balance to be used for individual trips. If a receiver was interested in the offer, they then had two weeks to go online to order the card. The trial period started four weeks after the send out date. This allows even those individuals who had just bought a monthly card to make use of the offer. Initial interest was recorded by receivers logging into the transport agency's webpage with a code from the flyer. Since the code was personal, we can track which treatment group the individual had been assigned to. While technically feasible, but for privacy reasons not legal, we do not collect any information on the demographics or precise location of the individual. We thus only track cards, not individuals.

If no reply to the offer was received, an informationally identical reminder was mailed out two weeks later (Figure A.4). The reminder had the same code as the first flyer, so we can track whether the card was claimed online. If a reply was received the resident received their free public transport card, the JoJo card, together with another letter (see Figure A.5). The letter that came with the card explained that the card could be used in the region of Skåne and on all buses, trams and regional trains. As the mailings were considered "business-as-usual" the participants were not informed that they were taking part in an experiment.

The data provides us with four types of outcome variables (see Figure 1 for an overview). First, we measure whether a card was ordered in response to the offer (or in response to the reminder). Second, we measured how many cards were activated. And third, we track the number of cards used and trips taken over an eight months period. The free trial period ended after two weeks (four weeks for one group), so the cards needed to be recharged by the participants themselves to be used at a later point. Fourth, we collect data on whether the cards have been uploaded with allowance for individual trips or with a monthly card.

Our data provides us with information on each card ID and how often it was used per day. We have no information on where users went with the card. We also cannot match the cards to individual people and their demographics for privacy reasons.

During the summer, 15th of June to 15th of August, Skånetrafik offers a very cheap and

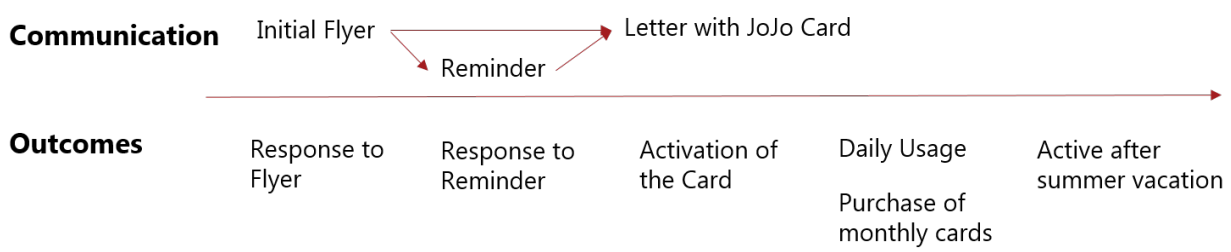

Figure 1: General process of the campaign and the outcome variables over time

popular summer card. While the design is colorful and it has the word "summer" on it, it works like the regular Jojo card, but it is only valid for that particular year and only the given time frame. The card costs 625 SEK (ca. USD 65) for those eight weeks. That makes it four times cheaper than two regular monthly cards. School-aged children receive a free summer card, and they can even take one of their parents and a sibling (under 20 years) along on a trip. The card is meant to encourage using public transport during summer when the demand is lower due to vacation time and to explore the region during school holidays. Given that three trips to neighboring cities in the region for one person would already cost the same as the summer card, the card is very popular. For our experiment, it means that there was a break in usage for most regular JoJo cards during the summer, as we will see in the data. We cannot track whether individuals bought the summer card instead. It does, however, provide the interesting opportunity to measure whether treatment differences remain after the 15th of August.

For our three experiments, we made use of the fact that the campaign runs in cohorts. For the first experiment, the names and addresses of everyone between the ages of 19 and 65 who moved to or within Skåne in the period between 12th of December 2017 and 16th of January 2018 were collected, which resulted in a sample of N = 14 282. These individuals received a randomized version of the first flyer on the 26th of February 2018. For the second experiment, we used people who had moved between the 26th of October 2017 and 11th of December 2017, who had received the control version of the first flyer and not responded. This group consisted of 11 373 individuals. For the third experiment, we contacted individuals who had responded to an initial untreated flyer between the 21st of February and 18th of March 2018 and who were now receiving their card in the mail. This group consisted of 6969 individuals. In total, 32,624 individuals were part of the three experiments, and no one was treated twice. We have data to track trips until the 29th of October 2018 for the cards in this first experiment. For experiments two and three we only have data for response and activation of the cards.

3 Main Experiment

In our main experiment, we test whether the addition of a social norm has an effect on the response to the offer of the free two-week JoJo card. Social norms can be injunctive (i.e., describing a socially accepted behavior) or descriptive (i.e. a norm that describes a prevalent behavior) (Cialdini 2003, Schultz et al. 2007). In their overview article Garcia-Sierra et al. (2015) propose the use of descriptive norms as a way to strengthen informational messages in promoting public transport. They argue that most pro-environmental behavior is driven by the desire to fit in with a social norm and that public transport usage can be modeled as a social dilemma. Driving a car might be optimal for the individual, but bad for society. Highlighting the social norm of public transport usage could affect beliefs about conditional cooperators and free-riding could be considered as a violation of the norm. Thus, if the majority uses public transport, this norm should be emphasized to discourage free-riding. Social norms have been proven effective in several other studies in the environmental domain. In one of the most prominent studies, Allcott (2011) evaluates a field experiment with a large US energy provider. Customers were shown their own energy consumption in comparison to their closest, most efficient neighbors. The intervention significantly reduced energy use. Similarly, letting people know that other hotel guests reuse their towels led to an increase in towel reuse (Cialdini 2003). In an overview article on green nudges, Carlsson et al. (forthcoming) present 17 field studies on the successful use of social norms in promoting environmentally friendly behavior. Fourteen of these studies are on reducing water or electricity consumption. Further, there is evidence of using social norms to encourage other socially desirable behavior such as paying taxes (Hallsworth et al. 2017) or giving to charity (Agerstrom et al. 2016). However, Donald et al. (2014) caution that descriptive norms might give an impression of public transport being overly congested and so discourage people from using it.

Since social proof needs to be honest, we used survey data from Skånetrafik to calculate the share of individuals who travel by public transport. The social norm message we used was: "Did you know that most people in Skåne travel with us? 72% of us travel by public transport occasionally. So join your neighbors and try it out". The number 72% comes from survey data that the transport agency had collected previously. To calculate 72% we aggregated the numbers of all individuals who travel at least several times per month to those who travel daily by public transport.

We test the social norm condition against a price instrument. We double the trial period to four weeks for the treatment group. The value of the four-week card is 1250 SEK (ca.

USD 130). Basic price theory tells us that this should make the offer more attractive. More importantly, though, there is evidence that it takes at least four weeks to form a habit. Given that public transport in Sweden is already quite popular (as we see from the social norm), we can assume that close to all individuals have previously experienced public transport. Thus, the campaign was targeted towards building a habit, rather than introducing individuals to public transportation as such. So while the higher value might attract more individuals, the longer time frame was an important aspect of this treatment.

We simply randomized our sample into three groups. The control group received the usual flyer with the offer to receive a free two-week travel card (N = 4 764). The social norms group receives the same two-week free travel card and the positive social norms message (N = 4 758). The four-week group receives the offer of a four-week free travel card without the social norm message (N = 4 760). See Appendix Figures A.1, A.2, A.3 for pictures of the flyers.

If individuals responded to the offer by ordering the card online, they received the standard letter with their card a couple of days later. If they did not reply, they received the standard reminder two weeks later. Otherwise, they received no further communication.

3.1 Results - Main Experiment

In this paper, we ask two primary questions. First, are nudges, compared to traditional price instruments, effective in changing the uptake and usage of public transport? Second, does increasing the trial period - from two to four weeks - lead to long-term habit formation? We will now present the results of the main experiment, which compares a social norm nudge to an increase in the travel period. For this sample, we also track long-term usage to measure habit formation.

Our three binary outcome measures are the initial response to the offer (entering the personal code online), the activation of the card during the trial period and the reactivation after the summer vacation, six months later. The two latter binary measures are equal to one if the card was held against a card reader in a bus, tram or train at least once during the trial period and after August 15th 2018, respectively. Because we are tracking cards, not individuals, there are no control variables to report.

Overall, we find that the average response rate to the offer is 44.6%. 84% of the cards that were claimed were also activated. That corresponds to an overall activation rate of around 38%. In the long-run 11% of the claimed cards were still in use (ca. 4% of the total offer).

We find that initial response to the free trial was higher in the four-week group (49.6 %, n = 2359) than in both the standard group (42.8 %, n = 2038) and the social norm group (41.6 %, n = 1977), (see Figure 2). A Chi2 test confirms that cards in the four-week version were significantly more likely to be claimed than in the standard group (Chi2 (1) = 44.03, $p < .001$). This means participants in the four-week group were 15.89% more likely to respond to the offer. Cards in the social norm group were no more likely to be claimed than those in the standard group (Chi2 (1) = 1.47, p = .225).

A larger percentage of cards was activated in the four week group (44.12 %, n = 2100) than in the standard group (34.45 %, n = 1641) and the social norm group (34.09 %, n = 1622). The difference between the activation rates is significantly different between the four-week group and the standard group, but not for the social norms group (Chi2 (1) = 93.38, $p < .001$ and Chi2 (1) = 0.13, p = .714, respectively).

After summer, we observe significantly more active cards in the sample of the initial four week offer (5.63%, n=268) than for the cards with the standard offer (3.76%, n=179) or the social norms offer (3.91%, n=186) (Chi2 (1) = 18.67,$p < .001$ and Chi2 (1) = 0.15, p = .700, respectively). At this time, users had had to upload money on their cards to continue using them.

Comparing the levels of the initial interest and the first activation with the usage after summer for all three groups, we can already conclude that the offer is popular and most individuals who claim the card also follow through with using the card at least once during the trial period. However, in the long-run, when users need to pay themselves for trips, only around 11% of the initially claimed cards are still in use. Clearly, there is a meaningful difference in behavior when travel is free and when it needs to be paid for.

The results from the non-parametric tests are confirmed in five probit regressions in Table 1. Since we have no control variables[3], we only regress our three binary outcome variables on the treatments. The standard two-week group is used as the baseline. For all regressions, we calculate robust standard errors. Column 1 shows the effects of the treatments on the initial response to the card. In the four-week treatment, the cards were 6.8 percentage points more likely to be claimed than in the two-week baseline. This effect is significant at the 1 percent level. The estimate for the social norms treatment is close to zero and not significant. Next we look at activation rates. Column 2 shows the activation rate of the cards. Cards in the four-week treatment were 9.7 percentage points more likely to be claimed than in the

[3]We know whether individuals who reacted to the campaign, live in rural, urban or suburban areas. We can, however, not match this to the card data used later in the analysis. We find no evidence of a differential selection effect based on location.

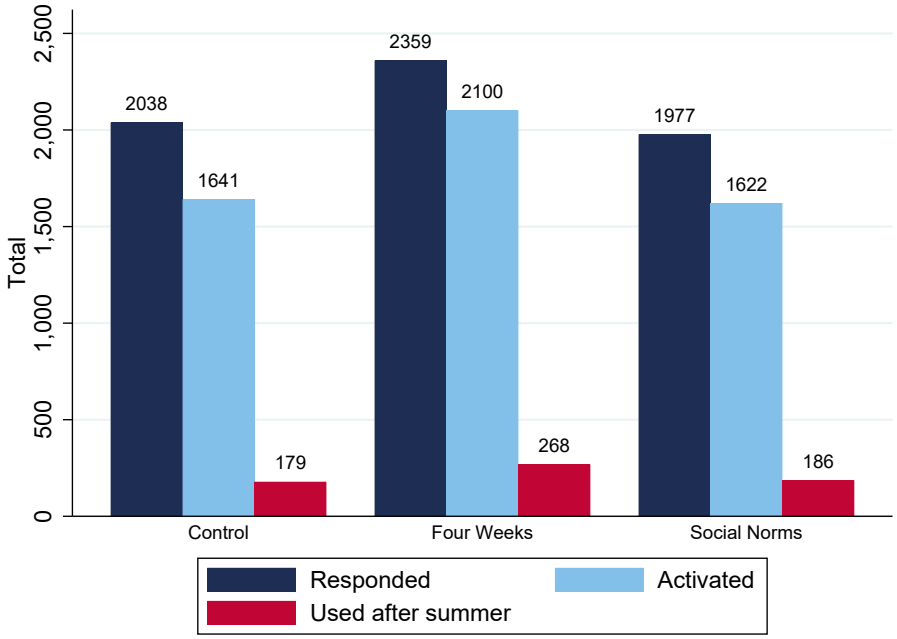

Figure 2: Number of responses, activated cards and cards used after summer in absolute numbers

baseline group. This effect is significant. The estimate for the social norms treatment is a clear zero. We do the same analysis based on activation after summer. Also for this outcome measure, we find a significant effect of the four-week treatment (Columns 3). The four-week treatment significantly increases the likelihood of a card being used after summer by 1.9 percentage points compared to the baseline. Unsurprisingly, there is no significant effect on the social norms message.

We can thus conclude this section with an answer to our first question. Our social norm nudge was not effective in changing travel behavior compared to the baseline offer. Increasing the economic incentives, on the other hand, was.

4 Supplementary Experiments 2 and 3

In our main experiment, we sampled from the full distribution of individuals who had moved within a certain time frame. Potentially, nudges could work better on samples who either did not respond to the pure economic incentive or who have shown their intention to act, but need an extra nudge to carry through with their action.

We thus conducted two further experiments on two types of sub-samples.

Table 1: Initial response, activiation of card and re-activation after summer

	Responded	Activated	After summer
Four Weeks	0.06772***	0.09663***	0.01892***
	(0.010)	(0.010)	(0.005)
Social Norms	-0.01240	-0.00367	0.00170
	(0.010)	(0.010)	(0.004)
N	14282	14282	14282
Pseudo R^2	0.004	0.007	0.005

Each column reports marginal effects from separate Probit regressions. Robust standard errors are in parentheses. There are no controls. The standard group is used as the baseline. Column 1 shows the effects of the treatments on the initial response to the card. Columns 2 and 3 show the unconditional estimates for activation of the card and usage after summer to show the intention to treat effect. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.1 Design Experiment 2

In our second experiment, we used a cohort of individuals who had not responded to the first offer of a two-week free travel card and was thus sent a reminder about the offer. These individuals have seen the same flyer as the individuals in the baseline group in the main experiment, but as they are drawn from a different cohort, they have not been treated before.

For this sample, we repeat the treatment comparison of a two-week card and a two-week card with the positively framed social norm used in the main experiment. We add a condition in which we frame the social norm negatively. Positively framed social norms are more common in published studies on social norms, but there is a discussion of whether negatively framed norms might be preferable in some settings. Levin et al. (1998) bases this idea on the seminal work by Kahneman & Tversky (1979), showing that people react stronger to losses than to gains. The same information framed as a loss compared to a reference point has been shown to lead to a stronger reaction than when framed as a gain. Avineri & Waygood (2013) explores positive or negative framing theoretically in relation to transport and environmental pollution, suggesting that negative frames might be more effective than positive ones in influencing choices of travel mode. We test this theoretical assumption in our experiment.

We had 11373 individuals available and randomized them into three equal groups. The first group (N = 3 791) received the basic reminder of the offer to travel for free for two

weeks. The second group (N = 3 791) received a reminder with the positive descriptive social norm "Did you know that most people in Skåne travel with us? 72% of us travel by public transport occasionally. So join your neighbors and try it out". The third group (n = 3 791) received the reminder with the negative social norm "Did you know that most people in Skåne travel with us? There are only 28% who do not travel by public transport. So join your neighbors and try it out".

If they responded to the offer, they then received the standard letter with their card a couple of days later. Otherwise, they received no further communication. See Appendix Figure A.4 for a picture of the flyers.

4.2 Results Experiment 2

For Experiment 2, our outcome measures are response to the reminder and activation of the card. We find no difference in response rates to the reminder between the standard group (20.7%, n=786), the positive norm group (21.0%, n=796) and the negative norm group (20.8%, n=789) (Chi2 test (2) = 0.08, p = 0.96). Further, there is no difference in the activation of cards. In the standard group 15.0% (n=572) of receivers activated their card, 15.3% (n=581) in the positive norm group and 14.7% (n=560) in the negative norm group (Chi2 test (2) = 0.729, p = 0.69).

The social norm nudge in either framing had no effect on response rates and activation.

4.3 Design Experiment 3

Even though people might have good intentions to change behavior, they tend to procrastinate and delay making decisions that are likely to be in their long-term interests (O'Donoghue & Rabin 1999). In the case of public transport usage compared to private car usage, individuals might intend to make more climate-friendly choices, but when it comes to act, time pressure, convenience, and weather conditions might derail their plans. A prompt to create a concrete plan for action, also known as an implementation intentions, has been shown to reduce the gap between intentions and actions, for example, for flu vaccination (Milkman et al. 2011), preventive screenings (Milkman et al. 2013), voting (Rogers et al. 2015) as well as for writing job applications (Abel et al. 2019). Implementation intentions work in three ways. First, merely asking the question of whether an individual is planning to engage in a particular behavior can increase the likelihood of the individual following through, as it brings the activity to the top of mind. Second, implementation intentions can help the

individual to make a plan and anticipate obstacles in advance. This could mean checking out the route to work in advance or planning which days a trip by public transport instead of by car would be feasible. Third, some implementation intentions are similar to promises, which would lead to internal discomfort if the plan is broken. While implementations have been shown to work in settings with one time behaviors, there is very limited evidence on implementation intentions on repeated behaviors. Carrera et al. (2018) provide gym-goers with prompts to schedule to go to the gym repeatedly and find no difference in exercise behavior compared to a group with no prompts. We test the hypothesis that implementation intentions increase the usage of free travel cards.

In this third experiment, we compare receiving a letter with instructions on how to use the card with two versions of implementation intentions.

We use a cohort of participants who had responded to the initial baseline offer of a two-week free travel card (N = 6969). So compared to the other two samples, these individuals have indicated their interest in receiving the card with the two week offer. We randomized them into three equal-sized groups of 2323 individuals each. Group 1 received the baseline letter with their card. In addition to the basic information on how to use the card, group 2's letter was extended by a statement saying that "based on experience, it is easier to start traveling when one makes a plan." The letter prompts the reader to think about when and where they will travel for the first time. Group 3 has the same text as group 2, but in addition, it has empty slots to fill out when and where they will travel to. It was not necessary to write down their plan. This treatment was inspired by Milkman et al. (2011), who show that for vaccination decisions, precise prompts that triggered thoughts about time day, time and place were the most effective. See Appendix Figure A.5 for a picture of the letters.

4.4 Results Experiment 3

For this third experiment, we measure activation of the card after receiving it in the mail. In the standard group 79.0% (n=1835) activated their card, 79.3% (n=1843) in the implementations intentions group without lines to fill in and 78.3% (n=1819) in the implementations intentions group with lines to fill in. The differences are not significant (Chi2 test (2) = 0.77, p = 0.680).

We conclude that the implementation intentions nudge had no effect on closing the intention action gap between ordering the free card and making at least one trip using the free card.

Our three experiments provide evidence that none of the popular behavioral interventions had an effect on response or activation rates of the cards. Neither for the full sample, nor for selected-sub samples. The economic incentives did increase response and activation rates, even after summer vacation.

4.5 Travel Behavior

Now that we have seen that receivers of the four-week trial offer in our main experiment were significantly more likely to reply to the offer and activate their card, we are interested in whether the treatments had an effect on travel behavior. we conduct the following analysis only on the sample of the main experiment.[4]

Figure 3 shows the aggregate number of trips taken over the eight months period (26th of March to 29th of October). The black line shows the two-week group and the red line the norms group, who are almost indistinguishable from each other. The blue line tracks the aggregate number of trips taken in the four-week group. The pattern shows a higher number of trips from Monday-Friday - consistent with a commuting pattern. This gives us confidence that the card is not only used for trips "on top" of necessary trips but is replacing other means of transport[5]. We also see a sharp drop for all three groups when the free trial period ends. Further, there are close to no trips during the summer months and then the trips pick up again after August 15th. The drop off around 15th of June is not as clear cut, as some people might have monthly cards that they use until they expire and then buy a summer card. For all measured time periods, the four-week group takes more trips per day than the baseline and the social norms group.

In Table 2, we present OLS regressions on the number of trips taken during different time periods to estimate the effect size of the differences and test for significance. We run the regressions on the full sample of experiment 1. The first column reports the effect for the first two weeks of the experiment. We use two weeks, so we can compare the number of trips taken during the experiment for the two two-week groups and the four-week group. We find that individuals assigned to receive the four-week card take 1.7 more trips during the first two weeks of the experiment than those in the baseline group. This effect is significant on the 1 percent level. There is no difference between the baseline and the social norms group. We see from the constant that in the baseline group the card is used for 5.9 trips during the

[4]We do not have follow up data for the second and third experiment, but given the clean null effects from the treatment there are no differences in travel behavior to be expected.

[5]We acknowledge, that we do not know which means of transport are being replaced — more on this in the discussion.

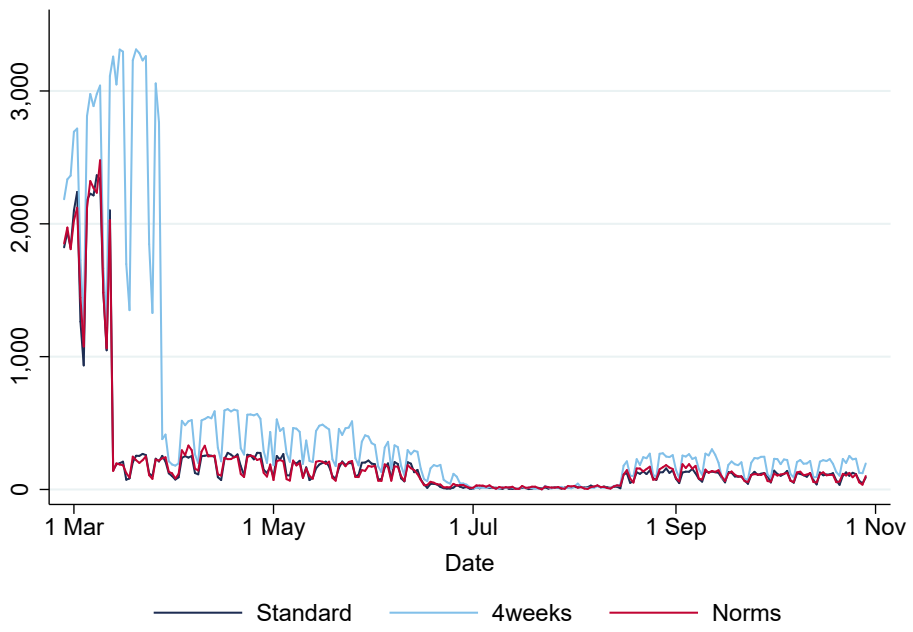

Figure 3: Number of trips taken per day divided by treatment group

two-week period. During the total experimental period the cards in the four-week treatment were used for 9.8 trips more than in the baseline (column 2). Given that these individuals had twice the time to use the card, it is not surprising that this is significantly more. The same mechanical increase is true for our regression in column 3 which shows the effect of the four-week treatment on the total number of trips ($\beta = 14.6$). However, excluding the experimental period to make the treatments comparable shows that cards in the four-week condition were used for on average 4.9 trips more than the cards in the two-week condition. Thus, the four-week treatment almost doubles the number of trips taken after the experiment is over (10.3 trips vs. 5.4 trips). None of the estimates for the social norms treatment are significant.

On average, individuals in the four-week treatment took 25.9 trips during the eight month period. This might seem quite low. It is important to remember, that these are unconditional results. More than 50% of the cards have not been claimed and thus have not been used for any trips. Further, as can be seen from Figure 3, there is close to zero usage of the cards during the summer period of 15th of June to 15th of August, because of the competing summer offer.

There is quite a sharp drop in usage after the offer ends. We can only speculate on the reasons. The most straight forward explanation is that for many individuals who own a

Table 2: Number of trips taken

	First 2 weeks	Experiment	Total	Total excl. Exp
Four Weeks	1.663***	9.755***	14.635***	4.880***
	(0.268)	(0.431)	(1.049)	(0.815)
Social Norms	0.055	0.055	0.175	0.120
	(0.251)	(0.251)	(0.805)	(0.687)
Constant	5.892***	5.892***	11.279***	5.386***
	(0.178)	(0.178)	(0.570)	(0.489)
N	14282	14282	14282	14282
R^2	0.004	0.057	0.020	0.004

Each column reports coefficients from separate OLS regressions using the full sample of Experiment 1. Robust standard errors are in parentheses. There are no controls. The standard group is used as the baseline. Column 1 shows the estimates for the effect of the treatments on the number of trips during the first two weeks of the experiment. Column 2 shows the effects on the number of trips during the experimental period (either for two or four weeks). Column 3 shows the estimates for the trips taken in the full eight months of our observation period. Finally, column 4 shows the estimates for the number of trips taken during the eight months of observation, excluding the trips taken during the experiment. All regressions are run on the original full sample (N=14282). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

car, free public transport is more cost-effective, but once they are asked to pay themselves, they go back to using their car. Other individuals might already have a JoJo card and go back to using their old one. Skånetrafik also offers an app with which one can pay for trips. Some individuals might decide to use the app instead. We cannot track transition from card to app. However, given that we have randomized the sample, these characteristics should only affect the level, not the treatment effect. In the next section, we will further explore whether the uptake of the card was purely a substitution with an old card among individuals who have already been taking public transport or also attracted and encouraged new travel habits among new public transport users.

4.6 Evidence of Habit Formation

The second question we are interested in answering is whether extending the trial period to one month, has a significant effect on habit formation. Habits are formed by repeatedly engaging in a behavior, with a more consistent performance of behavior leading to a habit being formed quicker (Lally et al. 2010). A longer incentivized time period increases the number of times a behavior can and will be performed, which leads to a higher chance of that behavior becoming a habit.

Some of the users of the free card might already be taking public transport regularly and thus have no more room to build a habit. We are interested in the group of individuals who has room to develop a habit, i.e. increase their usage of the card. As with most field experiments, we only observe the travel behavior of the individuals who chose to respond to the offer of the free travel card. Naturally, one might be concerned about self-selection based on unobservables. Even though both the two-week and the four-week treatment provide a significant incentive, especially to those individuals who already plan on commuting, we cannot rule out that the two offers attracted different types of individuals. As explained in section 2, the campaign was timed to make both cards attractive regardless of already purchased cards, by creating a time gap between the offer and the start of the trial period. That meant that even if someone had just bought a four-week card they were informed of the two-week offer early enough and could plan to use the free card for two weeks, when their current monthly card expired and before buying another monthly card. Monthly cards are not restricted to calendar days, but are valid for 30 days from day of purchase. Nevertheless, the unconditional higher trip rate in the first two weeks of the experiment points towards the groups not being completely equal.

In the following analysis, we can only look at activated cards. This means we cannot fully make use of the experimental design and clean randomization and the results should be interpreted with some caution. Nevertheless, investigating the travel patterns of the individuals who have opted in to each treatment can give us an insight into whether new habits were created during the experiment.

We split the data into activated cards that follow a commuting pattern of ten trips or more in the first week. Ten trips correspond to two trips per day - to and from work during workdays. We call these observations "high users" (N=1899). They hit the ground running from day one and might either be particularly motivated to change their habits or have already established commuting habits previously to the campaign. Since these users are already at a reasonable maximum of public transport usage, we cannot measure whether they are developing a new habit or not. A Chi2-test does not show any significant difference in "high users" active in the four-week or the two-week group (Chi2(1) = 0.863, p=0.353).

The interesting group for potential habit formation are the individuals who conduct less trips during the first week. Those users who use the card for less than 10 trips in the first week are classified as "low users" (N=3463). These individuals make up 65% of the group of individuals who have activated their card. They have the potential to increase their usage over time and do not show initial behavior that is in line with commuting and thus an

existing habit.

As we show in Figure 4, looking only at the high and low users of the four-week card, the "high users" by definition start high and continue high for the duration of the experiment and after (purple line). The "low users" (green line) increase their usage over the first three weeks of the experiment and level out during the last week and for the next eight months. We see the leveling out at a higher level then at the start as evidence that for these "low users" habit formation takes place and that a longer trial period leads to stronger habit formation. For example, the average number of trips in the first week of a low user in the two-week treatment is 3.6, while it is 3.2 in the four-week group. However, in the week after the experiment, the average number of trips per card of those classified as low users in the two-week group drops to 0.27, while it is still 1.3 for the four-week group. As they both started out with a similar number of trips (the four-week group even a bit lower), the post-experimental difference is unlikely to stem from initial differences in the users.

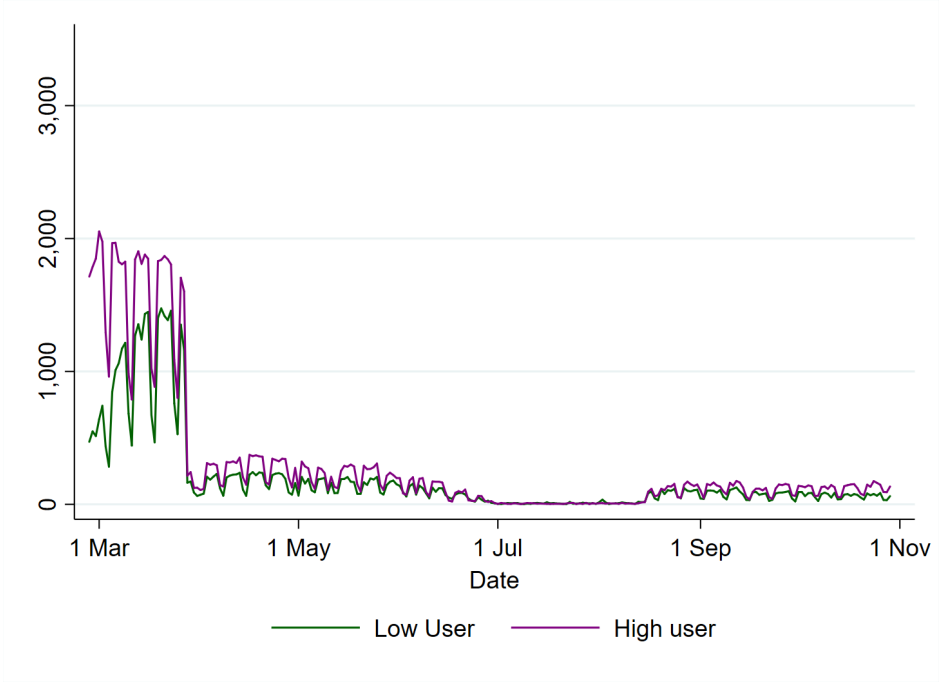

Figure 4: Total number of trips taken by high or low users

In Table 3 we show our results from an OLS regression with robust standard errors in which we regress the total number of trips taken during eight month, excluding the experimental period, on the treatments, the dummy of whether a card was used by a high user during the first week, and interaction effects of the treatments and the high user dummy. We cannot include the experimental period since the four week group had more time to take

free trips. From the perspective of habit formation, we are also more interested in effects after the incentivized period is over. We only include observations that have at least taken one trip during the experimental period.

We find a significant effect of the four-week treatment for the low users ($\beta = 6.383$). This means that in the eight months following the experiment, low users from the treatment group use the card, on average, 6.4 trips more than low users from the baseline group. This effect is significant on the 1 percent level. Unsurprisingly, those individuals we classified as high users also use the card more after the experiment is over. These cards are used for an average of 22 trips more than the cards of the low users. This is also significant on the 1 percent level.[6] There are no significant interaction effects of the high users and the treatments. Again, there is no significant effect on the social norms treatment. While these effects might seem small given the time frame, one must remember that there is a large share of zeros in the data, as many cards were not used after the experimental period. These effects are identified from ca. 31.7 percent of the initial responses.

The results in this section show that there is significant heterogeneity in the possibility of habit formation. Many individuals who respond to the offer behave in a way that is in line with existing habits. However, 65% of the users start from a low usage during the first week and increase over the four-week period. These individuals, who were randomized into either two or four-week cards seem to be positively affected by the additional two weeks, driving the treatment differences eight months later. The pattern we observe is thus in line with habit formation for this subset of individuals.[7]

4.7 Economic viability of the campaign

After the experimental period was over, the cards could be topped-up with either a balance that could be used for individual trips or with a monthly travel card. Table 4, columns 1 and 2, shows the amount of monthly cards bought for the sample of activated cards in the main experiment (n=5363). Cards in the four-week treatment were significantly more often uploaded with a monthly card ($p < 0.001$). On average, these cards were uploaded with 0.23 more monthly cards. Further, high users uploaded their card with 0.24 more monthly

[6] As a robustness check, we repeat the analysis with 20 trips the first two weeks and find comparable, significant estimates. In that estimation, the interaction term between four weeks and high users becomes significant as well. Results available by request.

[7] One might be concerned with a regression to the mean effect based on the initial classification of high and low users. We see that those individuals, who were initially classified as high users, stay above the ones classified as low users for all observation periods, but cannot completely rule out that there is some mean reversion in effect.

Table 3: Total number of trips taken based on initial travel pattern

	Total ex. Exp
Four Weeks	6.383***
	(1.593)
Social Norms	0.657
	0.657
High User	21.688***
	(3.491)
High User * Four Weeks	4.579
	(5.000)
High User * Social Norms	-1.032
	(4.802)
Constant	7.906***
	(1.028)
N	5363
R^2	0.038

The table reports the results of an OLS regression with the restricted sample of activated card from Experiment 1. The outcome variable is number of trips taken during the eight months of observation, without the trips taken during the free trial. Robust standard errors are in parentheses. The standard two-week group is used as the baseline. High users are individuals that have taken at least 10 trips during the first week of the free trial. The regression is run on the selected sample of cards that had been activated during the experiment (N=5363). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

cards. This is also significant. The estimate of the four-week treatment is not affected by controlling for high users. This lack of variation is additional evidence that the effect is not only driven by a different share of high users, but by the treatment.

Columns 3 and 4 show the total amount of money uploaded onto each card in the eight months period after the experiment. This can be in the form of monthly cards or pre-paid allowance. Cards in the four-week group were uploaded with on average 522 SEK, that is 201 SEK more than cards in the two-week group. Those classified as high users spend 145 SEK more.

Table 4: Number of monthly cards bought and total spent

	Monthly Card	Monthly Card	Total SEK	Total SEK
Four Weeks	0.230***	0.234***	201.417***	203.536***
	(0.035)	(0.034)	(34.960)	(34.857)
Social Norms	0.016	0.014	3.500	1.908
	(0.031)	(0.031)	(31.631)	(31.593)
High User		0.240***		145.273***
		(0.033)		(30.733)
Constant	0.250***	0.165***	321.358***	269.570***
	(0.021)	(0.022)	(22.479)	(23.341)
N	5363	5363	5363	5363
R^2	0.011	0.022	0.009	0.013

Each column reports coefficients from separate OLS regressions. Robust standard errors are in parentheses. The standard group is used as the baseline. Column 1 shows the estimates for the effect of the treatments on the number of monthly cards bought in the eight months after the trial. Column 2 includes a "high user" dummy. Columns 3 and 4 show the estimates for total spending. This could either be by buying a monthly card or by pre-paying the card for individual trips. All regressions are run on the selected sample of cards that had been activated (N=5363). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The cost of additional riders during the campaign period to the public transport agency is negligible. There were no additional direct costs. A quick back of the envelope calculation concludes that had everyone in the sample received the four-week card the transport agency would have made 107,000 SEK (10,900 USD) more. To calculate this we assume that also the two two-week groups would have 268 individuals active after summer and each of them would have spent 201 SEK more.

We can thus conclude that doubling the free trial period had a significant positive impact on habit formation for individuals who were not already using public transport to commute regularly.

5 Discussion

5.1 Survey data

Skånetrafiken conducted an online survey among the participants of our main experiment after the experimental period. Naturally, a selected sample opted-in to answer the survey (N= 257). Out of the individuals who had received a two-week travel card (either with the social norm or without) ca. 60% responded that they used the card 10 times or more during the trial period compared to the 85% in the four-week group, who gave this answer (see Figure A.6). 12% in the two-week group responded that they travel "much more" and 28% "a bit more" with public transport after the trial than before the trial period while for the four-week group it was 20% and 34% respectively (see Figure A.7). Our behavioral data is thus in line with this self-reported travel behavior, but should be considered with caution given the most likely strong self-selection into answering.

5.2 Salience and relevance of the social norm message

A valid concern could be that the social norm message wasn't salient enough compared to the four weeks of treatment. We did our best to restrict the information on the flyer and embed the social norm message in a way that would make it necessary to read the text to order the card. If the flyer was thrown out before the message was seen, then indeed, it might have been overlooked. We argue that this would be a bigger concern had we implemented a treatment with no free trial. An offer of two free weeks (65 USD) should have been of interest to many recipients. We leave it up to the reader to make up their mind whether this was salient enough. Further, while this argument is valid for the initial reaction to the card, it does not explain the effects on activation after receiving the card or the more intense use of the card. In our second experiment, we keep the monetary incentive constant and we still do not see an effect of the social norm message. Another concern could be the initial beliefs of the recipients of the offer. If they did read the social norm message, the honest statement of "72% of their neighbors traveling with public transport" might not have come as a surprise and affected their beliefs. While we could have attempted to test this by eliciting social norms prior to the experiment, we should consider what the practical outcome of that would have been. Given that 72% is the correct number and we are not willing to deceive in our experiments (and should not deceive when testing and implementing policy), we could not have sent them a false higher norm on the flyer, even if that might have affected the behavior of some. Further, eliciting beliefs prior to sending the flyers would have possibly

contaminated the findings of the natural field experiment. The experiment tested nudges in a very realistic and replicable setting. So even if, in principle, it could be possible to perfectly target and display the norm to some individuals, it is questionable how this should work in practice. Nevertheless, we acknowledge that our experiments test the effect of a social norm message in real-world marketing campaign and not the effect of social norms as such.

5.3 Substitution of means of transport

Because it is impossible to collect behavioral data on all possible means of transport (car, bus, bike, and walking), we cannot be sure whether the campaign encouraged the intended transition from car to public transport. From an organizational revenue perspective, it does not matter from which means of transport individuals were diverted, but from an environmental perspective, a transition from bike to bus or train is not desirable. We cannot rule out that for some individuals the card made public transport attractive compared to biking. Good weather will encourage public transport for those who would have taken their car, as they might need to supplement the bus or train with walking or biking for shorter distances. It will also encourage more people to bike rather than take public transport, given it is more pleasant to bike when the weather is nice (Liu et al. 2016). So it is unclear how the weather effect would show up in the data if we do not know the proportions of people's outside options. Other studies based on survey data provide evidence that free bus cards have encouraged a transition from car to public transport rather than from bike to public transport (Cooper 2007, Taniguchi & Fujii 2007, Thøgersen 2009). A commuting survey from Sweden (Trafa 2011) found that in the Malmö area, the largest city in Skåne, 58% of people living outside of Malmö, but working in Malmö, commute by car, while 40% commute by bus or train and only 2% commute by bike or walking. Even within the city 60% of commuters use the car, even though arguably, public transport infrastructure should be better within a city. Consequently, there is enough scope to switch from car to public transport and especially for the rural areas, the uptake cannot be explained by a switch away from bikes.

Figure 3 shows a steep decline in the usage of free travel cards after the free period is over. While the cards in the four-week treatment are more likely to be used after the experiment, overall, most users revert back to their travel habits before the offer. For some users this might mean using an old JoJo card. However, since the free cards and the cards that can be bought are identical, and they need to be recharged every month to be used as a monthly card, there is no obvious reason why individuals would want to go back to their old card

instead of continuing to use the new one. As stated before, they might switch to the phone app out of convenience. More research is needed to identify all substitution effects.

5.4 Self-Selection

Compared to some transportation demand studies that evaluate changes in travel behavior in response to exogenous shocks in prices or changes in infrastructure (Domencich & McFadden 1975, Larcom et al. 2017), our data only shows travel behavior of individuals who have opted in to the offer of receiving a free transport card. Clearly, this design choice create self-selection issues when looking at travel behavior. We have aimed to deal with these issues by using the full, unselected sample for most of our analysis, thus exploiting our clean randomization. Only in section 4.6 we need to use the selected sample of activated cards to explore travel habits among high and low users. Our study can thus speak to the question of whether different offers or nudges can generate interest in ordering a free travel card and using it. Randomized controlled trials like the ones presented in this paper can complement other transportation demand studies which usually have the weakness that shocks might not be completely exogenous and independent of other changes. In comparison to other habit formation studies such as Gneezy et al. (2011), John et al. (2011) or Royer et al. (2015), our challenges of a selection into the incentive scheme are at least identical or less problematic, as we can be sure that our sample was not aware that they were taking part in an experiment, while students in gym studies or employees in health studies are usually aware that they are being observed.

6 Conclusion

Using behavioral interventions such as nudges has been suggested as a viable way to change individual travel behavior. However, the evidence on whether public transport usage can be affected through non-monetary and non-infrastructure means is sparse. We show in three field experiments with a public transport operator that some of the most popular nudges that have been used successfully in other domains have no impact on travel behavior in our setting. A prolonged economic incentive does affect travel habits for a sub-set of individuals, even long after the incentive has been removed. Our findings are thus in line with the limited literature comparing economic incentives against nudges such as Ito (2015) who find that economic incentives have a longer-term effect on reducing energy consumption than a moral appeal. We find no evidence of social norm messages, framed either in a positive

or negative way on response rates or activation of free travel cards. This is in contrast to a large literature on social norms in the environmental domain. We further find no effect of implementation intentions on travel behavior. This finding is in line with Carrera et al. (2018), who do not find any evidence of implementation intentions affecting gym attendance, but in contrast to (Milkman et al. 2011, 2013) who find positive effects of a very similar intervention on vaccination decisions. While initial interest in the free travel cards is high, almost every second recipient orders the travel card, there is a steep decline in usage after the free period is over. From a policy perspective, the steep decline could be interpreted as evidence that public transport is experienced as too expensive for many individuals, compared to their preferred mode of transport. Providing free public transport in general or employers providing free travel cards to their employees might be a promising option in this setting, given that close to every second person in our representative sample of the Swedish population is at least willing to use public transport when it is available for free. Pure nudges, however, do not seem effective in overcoming the barriers to behavior change.

References

Abel, M., Burger, R., Carranza, E. & Piraino, P. (2019), 'Bridging the intention-behavior gap? the effect of plan-making prompts on job search and employment', *American Economic Journal: Applied Economics* **11**(2), 284–301.

Acland, D. & Levy, M. R. (2015), 'Naivete, projection bias, and habit formation in gym attendance', *Management Science* **61**(1), 146–160.

Agerstrom, J., Carlsson, R., Nicklasson, L. & Guntell, L. (2016), 'Using descriptive social norms to increase charitable giving: The power of local norms', *Journal of Economic Psychology* **52**, 147–153.

Allcott, H. (2011), 'Social norms and energy conservation', *Journal of Public Economics* **95**(9-10), 1082–1095.

Alta (2018), 'Behavioral insights to transportation demand management'.

Avineri, E. & Waygood, E. O. D. (2013), 'Applying valence framing to enhance the effect of information on transport-related carbon dioxide emissions', *Transportation Research Part A: Policy and Practice* **48**, 31–38.

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W. J. & Galing, S. (2017), 'Should governments invest more in nudging?', *Psychological Science* **28**(8), 1041–1055.

Carlsson, F., Gravert, C., Kurz, V. & Johannson-Stenman, O. (forthcoming), 'Nudging as an environmental policy instrument', *Review of Environmental Economics and Policy* .

Carrera, M., Royer, H., Stehr, M., Sydnor, J. & Taubinsky, D. (2018), 'The limits of simple implementation intentions: Evidence from a field experiment on making plans to exercise', *Journal of Health Economics* **62**, 95–104.

Cawley, J. & Price, J. A. (2013), 'A case study of a workplace wellness program that offers financial incentives for weight loss', *Journal of Health Economics* **32**(5), 794 – 803.

Charness, G. & Gneezy, U. (2009), 'Incentives to exercise', *Econometrica* **77**(3), 909–931.

Cialdini, R. B. (2003), 'Crafting normative messages to protect the environment', *Current directions in psychological science* **12**(4), 105–109.

Cooper, C. (2007), 'Successfully changing individual travel behavior: Applying community-based social marketing to travel choice', *Transportation Research Record* **2021**(1), 89–99.

Domencich, T. A. & McFadden, D. (1975), Urban travel demand-a behavioral analysis, Technical report.

Donald, I. J., Cooper, S. R. & Conchie, S. M. (2014), 'An extended theory of planned behaviour model of the psychological factors affecting commuters' transport mode use', *Journal of Environmental Psychology* **40**, 39–48.

European Commission, E. D.-G. (2019), *Guidelines for developing and implementing a Sustainable Urban Mobility Plan (2nd edition)*, Eltis.

Ferraro, P. J., Miranda, J. J. & Price, M. K. (2011), 'The persistence of treatment effects with norm-based policy instruments: evidence from a randomized environmental policy experiment', *American Economic Review* **101**(3), 318–22.

Finkelstein, E. A., Linnan, L. A., Tate, D. F. & Birken, B. E. (2007), 'A pilot study testing the effect of different levels of financial incentives on weight loss among overweight employees', *Journal of Occupational and Environmental Medicine* **49**(9), 981–989.

Garcia-Sierra, M., van den Bergh, J. & Miralles-Guasch, C. (2015), 'Behavioural economics, travel behaviour and environmental-transport policy', *Transportation Research Part D: Transport and Environment* **41**, 288–305.

Gneezy, U., Meier, S. & Rey-Biel, P. (2011), 'When and why incentives (don't) work to modify behavior', *Journal of Economic Perspectives* **25**(4), 191–210.

Goodwin, P. B. (1977), 'Habit and hysteresis in mode choice', *Urban studies* **14**(1), 95–98.

Gravert, C. & Kurz, V. (2019), 'Nudging à la carte: a field experiment on climate-friendly food choice', *Behavioural Public Policy* pp. 1–18.

Hallsworth, M., List, J. A., Metcalfe, R. D. & Vlaev, I. (2017), 'The behavioralist as tax collector: Using natural field experiments to enhance tax compliance', *Journal of Public Economics* **148**, 14–31.

IPCC (2014), 'Climate change 2014: Mitigation of climate change. contribution of working group iii to the fifth assessment report of the intergovernmental panel on climate change', *Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* **ISBN 978-1-107-05821-7**.

Ito, K. (2015), 'Asymmetric incentives in subsidies: Evidence from a large-scale electricity rebate program', *American Economic Journal: Economic Policy* **7**(3), 209–37.

John, L. K., Loewenstein, G., Troxel, A. B., Norton, L., Fassbender, J. E. & Volpp, K. G. (2011), 'Financial incentives for extended weight loss: a randomized, controlled trial', *Journal of general internal medicine* **26**(6), 621–626.

Just, D. R. & Price, J. (2013), 'Using incentives to encourage healthy eating in children', *Journal of Human Resources* **48**(4), 855–872.

Kahneman, D. & Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* **47**(2), 363–391.

Kallbekken, S. & Sælen, H. (2013), "nudging'hotel guests to reduce food waste as a win–win environmental measure', *Economics Letters* **119**(3), 325–327.

Kristal, A. S. & Whillans, A. V. (2019), 'What we can learn from five naturalistic field experiments that failed to shift commuter behaviour', *Nature Human Behaviour* pp. 1–8.

Kurz, V. (2018), 'Nudging to reduce meat consumption: Immediate and persistent effects of an intervention at a university restaurant', *Journal of Environmental Economics and Management* **90**, 317–341.

Lally, P., Van Jaarsveld, C. H., Potts, H. W. & Wardle, J. (2010), 'How are habits formed: Modelling habit formation in the real world', *European journal of social psychology* **40**(6), 998–1009.

Larcom, S., Rauch, F. & Willems, T. (2017), 'The Benefits of Forced Experimentation: Striking Evidence from the London Underground Network', *The Quarterly Journal of Economics* **132**(4), 2019–2055.

Levin, I. P., Schneider, S. L. & Gaeth, G. J. (1998), 'All frames are not created equal: A typology and critical analysis of framing effects', *Organizational behavior and human decision processes* **76**(2), 149–188.

Liu, C., Susilo, Y. O. & Karlström, A. (2016), 'Measuring the impacts of weather variability on home-based trip chaining behaviour: a focus on spatial heterogeneity', *Transportation* **43**(5), 843–867.

Metcalfe, R. & Dolan, P. (2012), 'Behavioural economics and its implications for transport', *Journal of Transport Geography* **24**, 503 – 511. Special Section on Theoretical Perspectives on Climate Change Mitigation in Transport.

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D. & Madrian, B. C. (2011), 'Using implementation intentions prompts to enhance influenza vaccination rates', *Proceedings of the National Academy of Sciences* **108**(26), 10415–10420.

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D. & Madrian, B. C. (2013), 'Planning prompts as a means of increasing preventive screening rates', *Preventive Medicine* **56**(1), 92–93.

O'Donoghue, T. & Rabin, M. (1999), 'Doing it now or later', *American Economic Review* **98**, 103–124.

Ortmann, A. & Dixit, V. (2017), 'Nudging towards a more efficient transportation system: A review of non-pricing (behavioural) interventions', *Unpublished Manuscript* .

Rogers, T., Milkman, K. L., John, L. K. & Norton, M. I. (2015), 'Beyond good intentions: Prompting people to make plans improves follow-through on important tasks', *Behavioral Science & Policy* **1**(2), 33–41.

Royer, H., Stehr, M. & Sydnor, J. (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.

Schaner, S. (2018), 'The persistent power of behavioral change: Long-run impacts of temporary savings subsidies for the poor', *American Economic Journal: Applied Economics* **10**(3), 67–100.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J. & Griskevicius, V. (2007), 'The constructive, destructive, and reconstructive power of social norms', *Psychological Science* **18**(5), 429–434.

Taniguchi, A. & Fujii, S. (2007), 'Promoting public transport using marketing techniques in mobility management and verifying their quantitative effects', *Transportation* **34**(1), 37.

Thaler, R. H. & Sunstein, C. R. (2009), *Nudge: Improving decisions about health, wealth, and happiness*, Penguin.

Thøgersen, J. (2009), 'Promoting public transport as a subscription service - effects of a free month travel card', *Transport Policy* **16**(6), 335–343.

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R. & Staake, T. (2016), 'Overcoming salience bias: How real-time feedback fosters resource conservation', *Management Science* **64**(3), 1458–1476.

Trafa (2011), Summary report: Commuting in stockholm, gothenburg and malmö – a current state analysis, Technical Report 2011:3, Transportanalys.

Verplanken, B., Walker, I., Davis, A. & Jurasek, M. (2008), 'Context change and travel mode choice: Combining the habit discontinuity and self-activation hypotheses', *Journal of Environmental Psychology* **28**(2), 121–127.

Volpp, K. G., Pauly, M. V., Loewenstein, G. & Bangsberg, D. (2009), 'P4p4p: an agenda for research on pay-for-performance for patients', *Health Affairs* **28**(1), 206–214.

Volpp, K. G., Troxel, A. B., Pauly, M. V., Glick, H. A., Puig, A., Asch, D. A., Galvin, R., Zhu, J., Wan, F., DeGuzman, J. et al. (2009), 'A randomized, controlled trial of financial incentives for smoking cessation', *New England Journal of Medicine* **360**(7), 699–709.

Appendix

Figure A.1: Initial flyer with the standard two weeks offer

Figure A.2: Initial flyer with the standard two weeks offer and social norms message

Figure A.3: Initial flyer with the four weeks offer

Figure A.4: Reminder after two weeks with positive social norm

Varsågod!

Nu börjar din resa

Som nyinflyttad har du blivit utvald att få pendla gratis med Skånetrafikens bussar och tåg i två veckor. Vi hoppas att du ska upptäcka hur enkelt och smidigt det är att resa med oss.

Ditt Jojo-kort gäller för obegränsat antal resor på våra stadsbussar regionbussar, Öresundståg, Pågatåg och Krösatåg inom hela Skåne, under angiven tidsperiod nedan. Värde 625 kronor.

Av erfarenhet vet vi att det är mer sannolikt att du kommer att använda ditt kort om du bestämmer dig för när och vart din första resa ska gå. Ta gärna och fundera igenom det redan nu. Vill du kan du också fylla i nedan för att göra det ännu enklare:

Jag ska åka från _____ till _____

_____ dagen den _____ /_____

När du reser håller du bara fram kortet nära och stilla mot kortläsaren på våra bussar tills den lyser grön. Du behöver inte säga vart du ska resa eftersom kortet gäller för resor i hela Skåne under din prova på-period. På tåget visar du upp kortet för tågvärden ombord.

Ditt Jojo-kort är klart att börja använda och gäller 30/4–14/5. Prova på-kortet gäller inte på Ventrafikens båtar.

Trevlig resa!

Figure A.5: Letter with implementation intentions

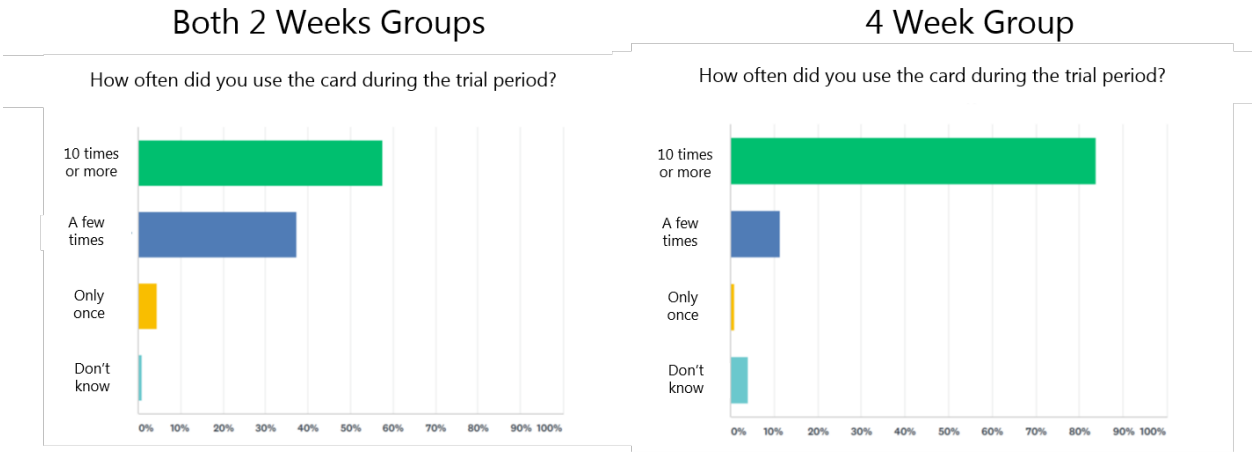

Figure A.6: Online Survey: How often did you travel with the card during the trial period? (N=257)

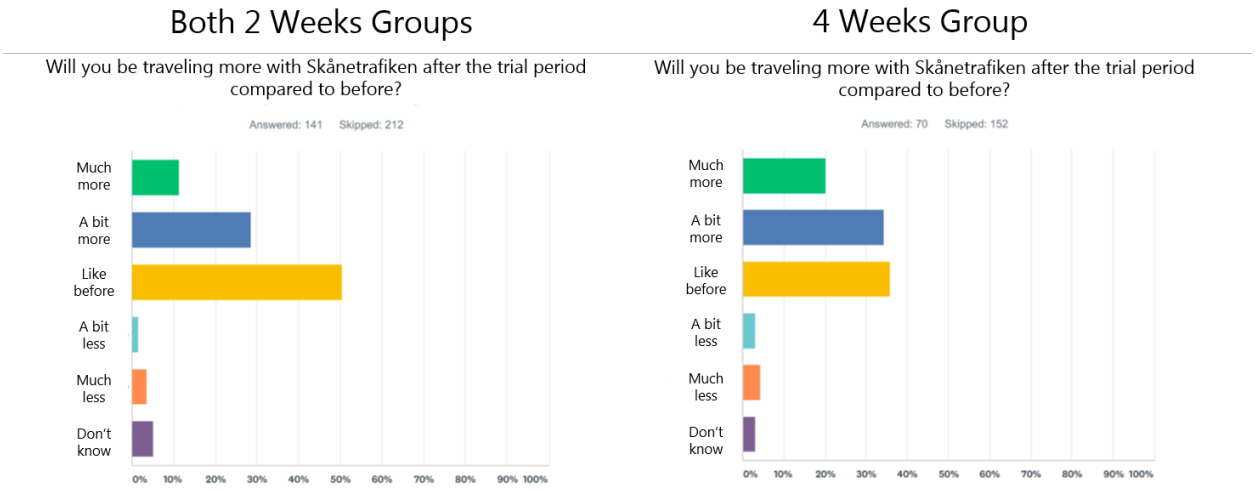

Figure A.7: Online Survey: Will you travel more with Skånetrafiken after the trial period than before? (N=221)