Transport Pricing to Promote E-biking and Reduce Externalities: Insights from a GPS-Tracked Experiment *

Jakob Roth† Laura Schwab† Beat Hintermann† Thomas Götschi†‡
Adrian Meister§ Lucas Meyer de Freitas§ Kay W. Axhausen§

April 28, 2025

Abstract

This study presents results from a randomized controlled trial involving 1,085 participants in Switzerland that have access to an E-bike, a car, and public transport. The participants' transport choices are monitored by means of a GPS-based tracking app. The treatment consists in a monetary incentive that approximates the main external costs and benefits associated with transport in the spirit of a Pigovian tax. This tax reduces transport-related external costs by 6.9 %, which corresponds to 78 Swiss francs per person and year (currently equivalent to 94 US dollars). The main underlying mechanism is a mode shift away from driving towards E-biking, public transport and walking. The results are primarily driven by individuals who own an S-pedelec with support up to 45 km/h, rather than users of the more common E-bikes that provide support up to 25 km/h. The pricing also induces a travel shift towards less congested time windows.

Keywords: Transport, Field experiment, GPS tracking, bicycle, E-bike, external costs, Pigovian taxation, transport pricing.

JEL Codes: H23, H31, I18, Q54, Q58, R41, R48.

*We thank the Swiss Federal Office of Energy and the cantons of Aargau, Basel-Landschaft, Basel-Stadt, and Zurich for funding.

†University of Basel, Peter Merian-Weg 6, 4002 Basel, Switzerland
‡University of Oregon, Hendricks Hall 119, Eugene, OR 97403, United States
§ETH Zurich, Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland

1 Introduction

The transition towards sustainable modes of transport has become an important goal for policymakers, given the environmental and economic significance of transport externalities. In the European Union (EU) and Switzerland, the monetized external costs amount to 1-2% of GDP (Van Essen et al., 2019; Bieler et al., 2019) and thus form a considerable burden for society. The main externalities in the transport sector include time loss due to congestion, as well as external costs and benefits related to health, accidents, and the environment (Van Essen et al., 2019). In contrast, maintenance costs are internalized through fuel taxes in many countries, including Switzerland.

Unlike private transport costs (e.g., fuel, public transport tickets), the external costs of transport are mostly ignored in people's private decisions about whether, when, and how to travel. The presence of external costs results in an inefficient use of the existing transport infrastructure in the sense of over-usage in terms of quantity and/or timing. This large-scale market failure presents a rationale for government action to implement policies that increase societal welfare (Eliasson, 2021). Following Pigou (1920) and Vickrey (1969), a tax amounting to the individuals' marginal external damage (Pigovian tax) is the most direct way to internalize these costs and, at least in principle, achieve an efficient outcome.

This paper extends the literature on the application of Pigovian taxation to the transport sector. We conduct a field experiment to examine how such a tax influences individuals' transport preferences and behaviors. Our sample consists of 1,085 Swiss residents who own an E-bike and regularly drive a car. We focus specifically on E-bikes as they are expected to be a better substitute for cars than conventional bicycles, given that E-bikes are faster and well-suited for longer distances. We employ a randomized controlled trial (RCT) design with a four-week baseline period followed by a five-week intervention phase. The study period spans September 2022 to July 2023 with rolling admission. Data on travel behavior and mode choice is gathered by detailed GPS tracking using a smartphone application. Pricing involves all main transport modes and is implemented by assigning personalized budgets to participants based on their baseline travel. External costs are then subtracted from these budgets, thus simulating a tax paid by participants.

In response to being charged the marginal external costs of transport, the participants in the treatment group reduce their external costs by 6.9%, starting from a baseline daily average of 3.35 CHF. This finding confirms the prior result found in Hintermann et al. (2024), who conducted a similar tracking study using a representative car-driving sample in Switzerland. We find that the treatment effect is mainly caused by a reduction in health- and congestion-related externalities. Regarding travel distances, the Pigovian tax leads to an average reduction of 8.2% (corresponding to 1.94 km) in daily car travel. At the same time, bicycle and walking distances increase by 12.6% and 6.1%, respectively, resulting in a significant shift away from driving towards active modes of transport. Public transport usage is also increased by 11.2%, while no effect on overall daily travel distance was found.

The effect is primarily driven by individuals who own an S-pedelec (i.e., an E-bike with support up to 45 km/h) rather than a regular E-bike. Mediation analysis shows that people not only reduce the amount of driving but also shift away from congested time windows. Overall, our results suggest that introducing the Pigovian rate reduces the external costs of transport through both mode and peak shifts.

To the extent that our pricing approximates the true societal costs of transport, this will necessarily translate into a societal net benefit. However, a Pigovian price is rarely implemented. We emphasize that our results also provide insights into the effects of other policy instruments that alter the relative prices of driving, public transport, and cycling. The further a policy instrument deviates from the Pigovian tax, the smaller the societal benefits will be, all else equal.

2 Related literature

A growing body of literature examines the potential for mode shifts towards E-bikes and the associated societal benefits, given their considerable rise in popularity. This subset of the broader mode choice literature focuses on the factors driving E-bike adoption, addressing both the extensive and intensive margins. Reviews of these determinants, including works by Fishman and Cherry (2016), Plazier et al. (2017), Bourne et al. (2020), and Buehler and Pucher (2021), highlight parallels with the factors that influence traditional bicycle use (Heinen et al., 2010). Among these, the built environment stands out as a key determinant influencing adoption (Smith et al., 2017).

At the extensive margin, i.e., the impact of owning or having access to an E-bike, several studies demonstrate a mode shift from car use to E-biking (Sun et al., 2020; Kroesen, 2017; Andersson et al., 2021). Bigazzi and Wong (2020) conduct a global review of stated preference surveys, revealing that E-bikes primarily substitute trips by public transport (33%), followed by conventional bicycles (27%), cars (24%), and walking (10%). However, these estimates heavily depend on the available infrastructure, and consequently, the predominantly substituted mode varies by study location. In the Swiss context, Reck et al. (2022) use revealed (GPS-based) preference data to estimate mode substitution for E-bikes and report that E-bikes predominantly replace car and public transport trips for distances greater than 2 km. Moser et al. (2018) analyze an intervention in Switzerland in which participants were offered an E-bike for a two-week trial in exchange for their car keys. They find that such habit-forming interventions have a long-term impact on mode choice. The environmental benefits of mode substitution towards E-bikes are further corroborated by Philips et al. (2022), McQueen et al. (2020), and Neves and Brand (2019).

Studies focusing exclusively on the intensive margin, that is, mode choice once an E-bike is already owned, are relatively scarce. While Heinen and Handy (2021) discuss various non-financial interventions in the real world, such as temporary car-free lanes, other studies analyze financial interventions to promote conventional cycling (Yang et al., 2010). To our knowledge, De Kruijf et al. (2018) is the only study to examine a financial incentive program specifically aimed at promoting E-biking. They find an increase in the share of commute trips made by E-bike instead of car, compared to a baseline period. However, the study design did not include a control group.

Since our study participants already own an E-bike, we focus on the intensive margin. In contrast to previous research, we price all transport modes based on their external costs. From an economic point of view, price-based instruments that accurately capture the external costs of transport are an efficient way to internalize externalities (Small and Verhoef, 2007; Verhoef, 2000). Real-world policies in this spirit are typically implemented in a second-best variant, such as fuel taxes (Santos, 2017; Charging, 2019), distance-related heavy vehicle fees (Krebs and Balmer, 2015), or congestion pricing schemes, typically in the form of cordon pricing. The latter accounts for both the location and the timing of travel, thus incorporating two key dimensions into the accurate price of the externality. Urban road congestion pricing has already been successfully introduced in several cities (Eliasson, 2021). Evidence from accompanying studies conducted in London (Leape, 2006; Transport for London, 2007), Milan (Gibson and Carnovale, 2015), Bergen (Isaksen and Johansen, 2021), Stockholm (Karlström and Franklin, 2009; Eliasson, 2009; Simeonova et al., 2021; Nilsson et al., 2024; Börjesson and Kristoffersson, 2018), Singapore (Agarwal and Koo, 2016; Olszewski and Xie, 2005), and Beijing (Yang et al., 2020) indicates that congestion charges are successful in reducing both traffic congestion and, for some, air pollution levels. These externality reductions are mediated by departure time shifts to off-peak times and mode shifts towards public transit, and, to a lesser extent, active modes. The first-best tax, in the spirit of Pigou (1920), which charges all transport modes for their marginal external costs based on intensity, time, and place, remains unimplemented.

Beyond implemented policies, several experiments assess the effects of dynamic transport pricing. Nielsen (2004) uses the GPS tracking of 500 cars in Copenhagen to study a congestion charge with two rates, finding that the higher rate induced shifts in mode and travel time. Ben-Elia and Ettema (2011) use in-vehicle tracking in the Netherlands to study the impact of various rewards on commuter behavior. They find that financial incentives reduce rush hour driving, encourage off-peak travel, and increase public transit use, cycling, and remote work. Martin and Thornton (2018) conduct a road pricing experiment with GPS tracking in 1,400 vehicles, testing various charging schemes. They find that constant charges reduce high-speed and off-peak road use, while peak-time or central-area charges are more effective in reducing congestion. Tsirimpa et al. (2019) show that reward-based instruments, which incentivize participants toward sustainable multimodal choices using a smartphone app, increase the use of public transport and walking. The "Traffic Choices Study" in Seattle finds that variable road tolling with GPS tracking leads to a significant reduction in congestion, as travelers adjust their routes, times, and modes of transport, thereby improving overall traffic efficiency (Council, 2008). "BART Perks", a six-month pilot program in San Francisco, offered cash rewards to encourage riders to shift their trips towards off-peak hours, resulting in a 10% reduction in peak-hour travel (Greene-Roesel et al., 2018). The Metropolitan Transport Commission's pilot program in San Francisco used GPS tracking via an app and offered \$3-\$5 financial incentives to encourage shifts to sustainable modes. The program effectively increased the use of intermodal transport, walking, cycling, and public transit, particularly when alternatives were accessible or appealing (Metropia, 2024).

Research on transport interventions often lacks a never-treated control group, making it difficult to isolate the treatment effect from other concurrent dynamics. RCTs address this by using a control group, randomly assigned, which is exposed to the same external factors and, on average, is identical to the treatment group. We are aware of several other RCTs in the transport sector that incorporate financial incentives. Rosenfield et al. (2020) conduct an RCT with both an informational campaign and monetary incentives involving 2,000 employees at the Massachusetts Institute of Technology aimed at reducing parking demand. They find no significant reduction in car usage in any of the three treatment arms, nor any increase in the use of alternative modes. Kreindler (2024) conducts a natural field experiment to examine the effect of peak-hour traffic congestion pricing in Bangalore using a smartphone app. The commuter responses reveal moderate schedule inflexibility and a high value of time. Hahn et al. (2024) estimate price elasticities for urban mass transit in San Francisco using a large natural experiment and a natural field experiment with price reductions during time windows with less crowding. They find that off-peak subsidies can increase welfare, but the effects diminish when consumers consider others' decisions. The most closely related experiment to our study is Hintermann et al. (2024), who implement both a Pigovian tax and an informational nudge in a representative car-driving sample in Switzerland using a smartphone tracking app. They find a significant reduction in transport-related external costs due to the pricing treatment, driven by mode substitution and shifts in departure times.

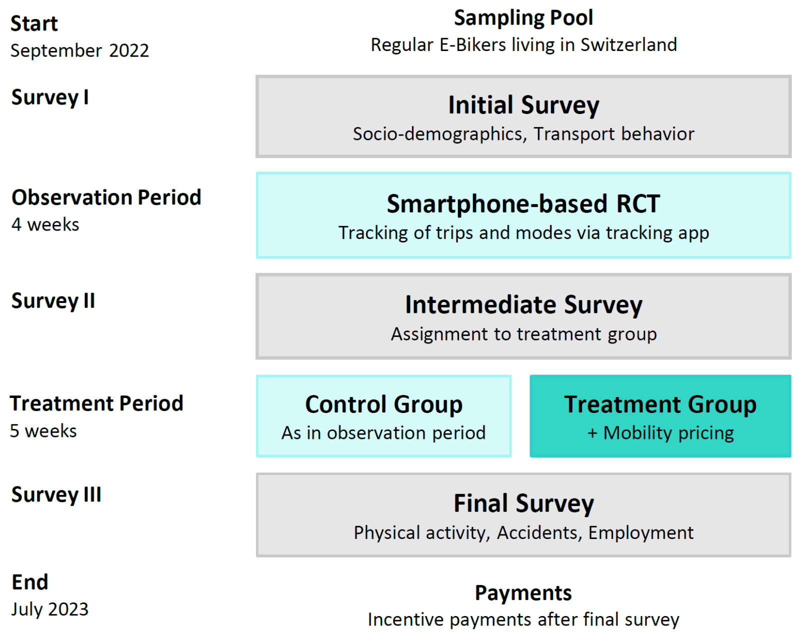
However, the study cannot identify a statistically significant increase in cycling, likely due to limited statistical power and the small number of cycle users. Previous studies clearly highlight the substitutability of car-driving with E-biking, which motivates an exploration of the effects of a first-best Pigovian tax on transport externalities among E-bike users. Our study follows a similar experimental setup to Hintermann et al. (2024), which we consider ideal for testing the effects we aim to identify. By focusing on E-bikers who still regularly drive, we can conduct a detailed heterogeneity analysis of this growing subgroup, who likely has the potential to replace car trips with E-bike trips.

3 Experimental setting

3.1 Study design and sample

To estimate the causal effect of charging marginal external costs, we endowed participants with a travel budget based on their baseline needs. The tax, which was an approximation of the external costs, was subsequently deducted from this budget. Participants received weekly mobility reports showing the external costs, the five costliest travels, and their budget balance. At the end of the study, they received the remainder of their budget on top of the 50 Swiss francs (CHF) study reward.[1] The study sample consisted of 1,085 E-bike users living in the German- and French-speaking parts of Switzerland. The study used rolling recruitment and lasted from September 15, 2022, to July 31, 2023. Two waves of participation emerged, with many people starting simultaneously: an initial wave from broad recruitment across multiple channels, followed by a second surge in February after invitations were sent via the Zurich vehicle registration office.[2]

Figure 1: RCT study design



In an initial online survey on travel behavior and demographics, participants were screened and invited into the RCT or another part of the research project (see Heinonen et al. (2024) for an overview of the entire project). To qualify for the RCT, respondents had to be at least 18 years old, live in Switzerland, own an E-bike,[3] use a car at least twice

[1] At the beginning of the study, the Swiss Franc was almost exactly at parity with the US dollar.

[2] We will analyze these two waves separately in subsection 5.3.

[3] In Switzerland, two types of E-bikes were in use during the study period: "regular" E-bikes (or "pedelecs"), which provide electric assistance up to 25 km/h, and "S-pedelecs", which assist up to 45 km/h. The latter require a helmet and registration as a motor vehicle. Owners of both E-bike types were invited to the experiment.

per week, and not use a regular bicycle.[4] Qualifying individuals were then invited to the tracking phase, which started with an observation period during which participants received a weekly summary of their travel behavior by e-mail. Then, participants were invited to fill out the intermediate survey, in which the treatment was explained and delivered. During the treatment period, participants in the treatment group received an extended weekly report containing information on their external costs. To receive their incentive payment, they needed to complete a final survey.

To cost effectively obtain a sufficient number of E-bike users, we used multiple, targeted recruitment channels. The largest group of participants was invited by e-mail, with 32.3% contacted through the research institute "YouGov"[5], and 20.6% via personalized e-mails sent to addresses provided by the cycling organization "Pro Velo"[6]. Another portion of the respondents were contacted by mail through cantonal vehicle registration offices (18.9%), which maintain records of S-pedelec owners. The remaining individuals were reached via invitations on the intranet of cantonal administrations (3.4%), social media posts on Instagram and Facebook (2.5%), newsletters of cyclist organizations (2.4%), or directly from our website (19.1%).[7] Heinonen et al. (2024) present the recruitment strategy and the response rates in more detail.

Table 1 summarizes key socio-demographic variables for both the introduction survey sample and the tracking sample (which is a subset of the former). We also present the corresponding variables from the Mobility and Transport Microcensus (MTMC) sample, a representative survey of Swiss travel habits conducted by the Federal Office of Statistics and the Federal Office of Spatial Development (2023). Comparing the eligible participants from the introduction survey to the RCT sample provides an indication of the selection bias introduced by the fact that not all individuals are willing to be tracked via a smartphone app. Specifically, individuals aged 66 to 87 years, with a secondary education level, and living in two-person households are less likely to agree to participate in the tracking study.

We limit the representative MTMC sample to individuals aged 18 to 87 with access to both a car and an E-bike to ensure a meaningful comparison with our sample. Overall, our recruitment strategy resulted in an RCT sample that closely mirrors the MTMC population in terms of observable characteristics, with some exceptions. For instance, the tracking sample includes fewer young adults (aged 18–30) and fewer females, while participants tend to have higher incomes. Our sample also has a higher share of S-pedelecs due to our recruitment strategy (see above). It shows a similar distribution of mode shares but

[4]The app cannot distinguish between bicycles and E-bikes. To ensure that we were observing E-bike trips, we limited the sample to individuals who rarely or never use a regular bicycle. Additionally, individuals were excluded from the RCT if they were professional drivers (e.g., taxi drivers, bicycle couriers, or train drivers), did not own a smartphone, or were unable to walk 200 meters without assistance.

[5]Formerly LINK institute, `https://business.yougov.com/`

[6]`https://en.pro-velo.ch/`

[7]The website's URL was provided in the cantonal letters and social media posts. Unfortunately, we cannot clearly determine the origin of these study entries.

Table 1: Demographic sample information

Variable	Level	Intro survey	RCT sample			MTMC
			Control	p-value	Treated	
Age	(18 - 30]	2.4	3.1	0.771	3.5	16.7
	(31 - 50]	29.1	42.1	0.957	42.3	34.2
	(51 - 65]	44.2	40.7	0.492	42.8	31.4
	(66 - 87]	24.3	14.0	0.211	11.4	17.7
Education	Mandatory	3.2	2.0	0.485	2.6	7.5
	Secondary	59.9	50.0	0.512	52.0	57.7
	Higher	36.8	48.0	0.327	45.1	34.8
Gender	Female	43.4	41.3	0.121	36.7	50.4
	Male	56.6	58.7	0.121	63.3	49.6
Household size	1	11.3	11.6	0.071	8.3	9.9
	2	49.6	38.6	0.494	36.6	39.4
	3	13.3	16.3	0.663	17.3	18.4
	4	18.9	24.8	0.402	27.0	22.3
	5 or more	7.0	8.7	0.245	10.7	9.9
Monthly household income	4,000 CHF or less	3.4	2.4	0.028*	0.7	5.3
	4,000 - 8,000 CHF	26.6	22.4	0.864	22.9	24.7
	8,000 - 12,000 CHF	23.9	25.6	0.12	29.8	25.6
	12,000 - 16,000 CHF	24.2	30.7	0.056	25.5	13.8
	16,000 CHF or more	12.3	14.2	0.114	17.7	11.6
	Prefer not to say	8.3	4.5	0.04*	2.3	8.3
	I don't know	1.3	0.2	0.041*	1.2	10.6
Language	German	85.4	81.9	0.039*	86.5	78.7
	French	12.0	14.4	0.073	10.7	18.0
	Italian					3.3
	English	2.6	3.7	0.372	2.8	
Nationality	Swiss	86.5	82.7	0.445	84.4	86.0
	Other	13.5	17.3	0.445	15.6	14.0
Residential setting	Rural	15.6	12.2	0.417	13.9	20.4
	Periurban	28.5	27.2	0.885	27.6	26.1
	Urban	55.9	60.6	0.492	58.6	53.6
Access to car	Yes	96.6	95.3	0.653	95.8	72.2
	Sometimes	3.4	4.7	0.653	4.2	22.7
	No		0.0		0.0	5.2
E-bike (25 km/h) ownership	Yes	63.9	59.6	0.596	58.1	89.5
E-bike (45 km/h) ownership	Yes	44.0	49.0	0.419	51.5	14.8
Full public transport subscription	Yes	6.4	7.7	0.487	8.8	14.3
Half fare public transport subscription	Yes	66.8	69.9	0.814	70.5	45.6
Distance	Car distance (km)		28.4	0.068	26.3	26.7
	Public transport distance (km)		9.1	0.226	10.2	6.7
	E-bike distance (km)		5.1	0.591	4.9	0.9
	Bicycle distance (km)					0.9
	Walking distance (km)		1.8	0.154	1.9	1.5
	Total distance (km)		44.7	0.412	43.6	40.2
Duration	Total duration (min)		92.1	0.774	91.5	84.2
External costs	Climate ext. costs (CHF)		1.1	0.136	1.1	
	Congestion ext. costs (CHF)		0.7	0.207	0.7	
	Health ext. benefits (CHF)		-1.1	0.853	-1.1	
	Health ext. costs (CHF)		1.2	0.264	1.2	
	Accident ext. costs (CHF)		1.5	0.305	1.4	
	Total ext. costs (CHF)		3.5	0.157	3.3	
Private costs	Private costs (CHF)		10.4	0.509	10.2	
Recruitment wave	Autumn	74.6	71.7	0.657	72.9	
N		5,993	508		577	11,176

Notes: Descriptive statistics for all individuals eligible for the study, the RCT study sample, and the comparable weighted sample from the Swiss Mobility and Transport Microcensus 2021 (MTMC), including households with at least one E-bike and one car. The first panel presents percentages, while the bottom panel shows baseline averages with units given in parentheses. * p-value < 0.05 indicates significant differences between the control and treatment groups, without correction for multiple hypothesis testing.

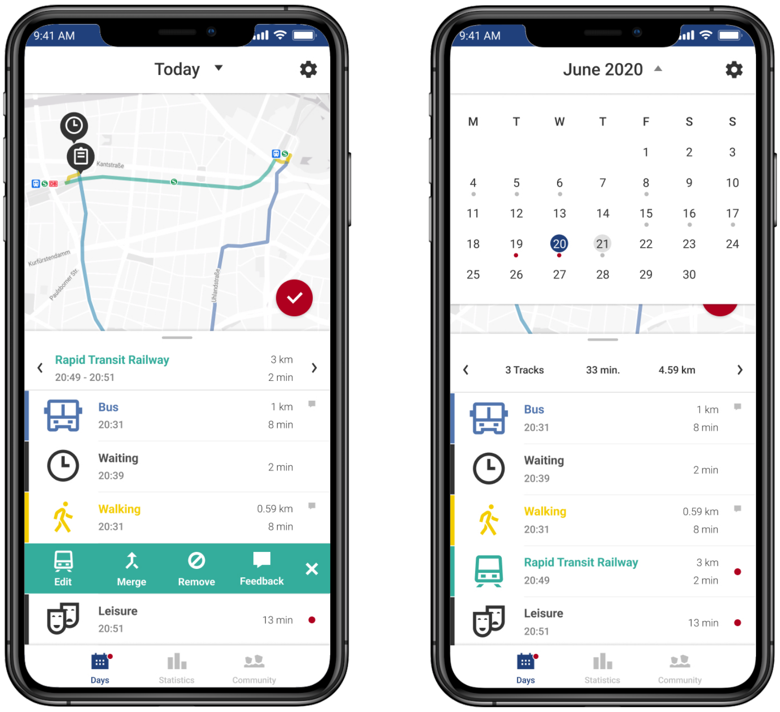
slightly higher overall travel frequency. Since individuals who regularly use a conventional bicycle were excluded from the RCT, no cycling distance is reported for this sample.

The randomization into treatment and control groups was effective, as indicated by the p-values from two-sample mean tests. Most variables are evenly distributed across both groups, particularly the pre-treatment averages of the main outcome variables, as shown in the lower part of Table 1. However, the treatment group includes a slightly higher proportion of German-speaking individuals. The table also highlights differences in some income responses, though these are based on very few observations, with only 16 individuals reporting an income below 4,000 CHF.[8]

3.2 GPS tracking

To collect data on participants' transport behavior, we deployed the smartphone tracking app "Catch my Day".[9] The app has been successfully employed in several previous transport studies (see e.g., Hintermann et al., 2024; Molloy et al., 2020). Figure 2 displays two screenshots of the tracking app.

Figure 2: The "Catch my Day" interface



Note: GPS tracking app on iPhone (left: map view, right: calendar view).

Once installed and activated, the app continuously records the location, direction of travel, and speed of the smartphone. The underlying technology automatically detects the

[8]Group assignment was random, meaning these differences occurred by chance. Balance in specific attributes could have been enforced through stratified randomization, but we opted against this due to the technical challenges of implementing it in a rolling admission setup.

[9]The app's developer MOTIONTAG is a specialist in data-driven services in the transport field. The app is publicly available for iOS and Android smartphones but requires an access code for use.

mode of transport with 92% accuracy using machine learning-based imputation methods (Molloy et al., 2020). Due to similar speeds and movement patterns, the app cannot distinguish between bicycles and their electrified counterparts. To address this, we required participants to ride a conventional bicycle no more than twice per week, allowing us to attribute all recorded bicycle trips to E-bikes with reasonable certainty.[10] We consider only trips categorized as car, motorcycle,[11] public transport (all train types, bus, tram), E-biking, and walking. We ignore all movements made using other modes of transport. Participants had the option to manually confirm or correct the detected mode and provide information about the activities between movements (i.e., the purpose of the trips). While allowing participants to review their tracks can improve accuracy, it also introduces the potential for experimental manipulation. For this reason, we compare our results using the automatically detected mode and the participants' corrections in subsection A.2.

3.3 The external costs of transport

To compute the marginal external costs of transport, we consider the categories of congestion, climate, health, and accidents. For each category, we focus solely on the external component of these costs, i.e., on the costs (or benefits) that accrue to society at large and are not paid for by the person who makes the transport decision. The most straightforward examples are climate damage and damage from noise and local air pollution, which essentially have no internal component.

Accident-related external costs predominantly consist of the inflicted health costs which are socialized through the insurance system in Switzerland. There is no deductible for accident-related health costs, such that some person A pays for none of the monetary costs out of pocket, rendering them entirely external. Accidents also cause staff shortages and costs for replacing injured or deceased persons. In contrast, damages to vehicles are excluded, as the damage to person B's car caused by person A will be paid either directly by person A or indirectly via a raise in A's personal car insurance premium. Similarly, health costs due to local air pollutants and noise emission are external to traffic participant A. We include these costs separately from the external health benefits of active transport, which are the savings in health care costs due to an increase in physical exercise. We stress that the reduction in mortality due to exercise is not included in our numbers, as this is an internal component (it is the cyclists and walkers themselves who live longer).

This approach and the corresponding external costs are based on official values published by the Swiss Federal Office for Spatial Development in 2021 (Bieler et al., 2019). However, this report lacks indications for electric cars and E-bikes. We therefore estimated these values based on their non-electric counterparts and the emission factors for

[10]Similarly, we assume all car trips are made by the type of car (combustion, hybrid, or electric) that the respondents declared as their main car in the initial survey.

[11]Due to the very small mode share of 0.3% of all trip stages for motorcycles in our sample, we omit the results for this mode in the rest of the paper.

electricity production in Switzerland (Sacchi and Bauer, 2023). Furthermore, we adjusted the accident-related external costs for E-biking such that they exactly offset the external benefits, resulting in net zero external costs for this mode.[12] In a recent update of the external costs calculations, the Federal Office for Spatial Development now estimates an external benefit of 2.1 cents per km for E-bikes (Ecoplan-INFRAS, 2024), which is very close to the adjusted rate calculated for this study.

The resulting values are shown in Table 2. For public transport, walking, and E-biking, the external costs are constant per kilometer, as differences across time and space are not considered. For walking, the health benefits outweigh the accident costs, such that this mode exhibits net external benefits.

Table 2: Marginal external costs by mode

		Car	E-Car	Motorcycle	E-Bike	Walking	Bus	Tram	Train
Climate & environment	Climate	1.52	0.09	1.05	-	-	0.76	-	0.01
	Nature and landscape	0.87	0.68	0.48	0.50	0.36	0.26	0.03	0.47
	(Toxic-) ground poll.	0.08	0.08	0.05	-	-	0.16	-	0.12
	Up-/downstream proc.	0.86	0.46	0.76	1.36	-	0.46	0.66	0.17
	Urbanisation/separation	0.22	0.22	0.16	-	-	0.17	0.13	0.15
	Total	3.55	1.52	2.51	1.86	0.36	1.81	0.82	0.92
Accidents		2.15	2.15	14.32	12.85	7.97	2.70	1.34	0.16
Health benefits		-	-	-	-14.72	-18.30	-	-	-
Health costs	Local pollutants	2.57	1.93	0.69	-	-	-	-	1.42
	Noise	1.04	0.78	14.77	-	-	1.01	0.15	0.86
	Total	3.61	2.71	15.46	-	-	1.01	0.15	2.29
Congestion	Average†	2.65	2.67	0.66	-	-	-	-	-
Total		**11.96**	**9.05**	**32.95**	**0.00**	**-9.97**	**5.51**	**2.31**	**3.37**

Notes: Values in Swiss cents per person-kilometer, based on Bieler et al. (2019) and own calculations for E-cars and E-bikes. † The congestion values are observed averages in the data (see below for congestion costs). For many trips, the congestion externality is zero.

For congestion, the external costs consist of the marginal time loss imposed on others as a result of choosing to participate in traffic at that time and location.

The external cost of congestion does *not* include the time loss from congestion incurred by the driver, as this is presumably internalized in their choice to drive. For cars and motorcycles, the external costs consist of a fixed per-km rate (which differs between regular and electric cars) and a time- and location-specific component for the external cost of congestion. However, providing the participants with a continuous menu of prices is not ideal. To simplify the price schedule and thus increase its salience, we discretize the continuous distribution of congestion costs by computing the average congestion externality for three urbanization levels and four different time periods per day (Table 3). This approach is

[12]This reduction was implemented because of the difficulty conveying positive net external costs of cycling to E-bikers. To achieve a net zero value, the external accident costs provided by the Swiss Federal Government have to be cut in half. This reduction leads to external cost numbers that are very close to those reported for the Netherlands or Denmark (Castro et al., 2018). These countries have lower bicycle-related accidents (per km) due to a better infrastructure and a higher bicycle mode share, leading to "safety in numbers". In this sense, our numbers reflect a future in which cycling in Switzerland is similar in terms of safety to cycling in these countries.

based on the external congestion costs observed in Hintermann et al. (2024) that were estimated using the approach described in Molloy et al. (2021). The degree of urbanization for both origin and destination points is based on the official classification provided by Eurostat (2021). The classification includes 1 for dense urban areas, 2 for medium-density areas, and 3 for sparsely populated areas. As shown in Table 3, the highest congestion costs occur during the evening peak within urban areas.

Table 3: Congestion costs of car travel for different times and regions

Trip Direction: (origin→destination)	**1→1**	**1→2**	**1→3**	**2→1**	**2→2**	**2→3**	**3→1**	**3→2**	**3→3**
Morning Rush-Hour 6:30-8:30	7.0	2.5	1.9	4.5	2.0	1.3	3.0	1.1	0.6
Off-peak Hours 8:30-16:30 & 18:30-20:00	5.7	3.3	2.4	3.0	1.9	1.0	2.1	1.1	0.5
Evening Rush-Hour 16:30-18:30	10.3	6.7	4.2	5.2	3.2	1.7	2.9	1.7	1.2
Night 20:00-6:30	0	0	0	0	0	0	0	0	0

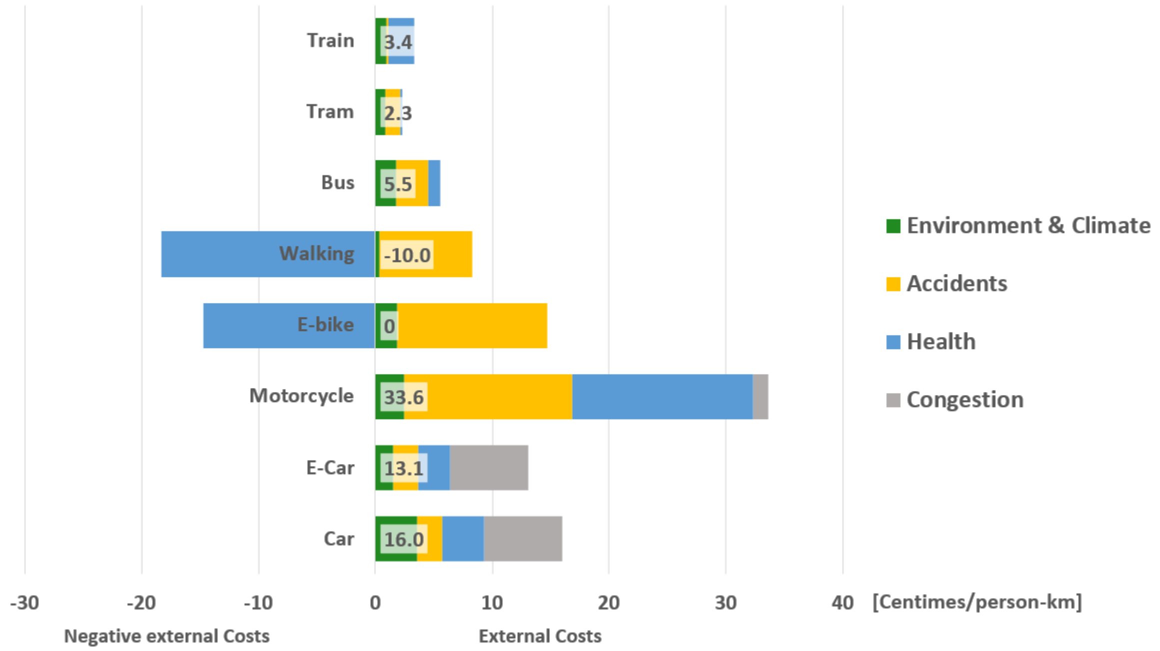
Notes: Congestion costs denoted in cents/km. 1: Dense urban area (cities), 2: Medium density area (towns and suburbs), 3: Sparsely populated area (rural areas). Reading example: A trip from an urban area (1) to a medium density area (2) during evening rush hour is assigned an external congestion cost of 6.7 cents per km.

3.4 Treatment

During the observation period, participants received a weekly summary of their travel behavior, including duration and distance traveled by mode (Figure 4a). After this period, all participants were invited to the intermediate survey, in the beginning of which the participants were randomly assigned to either the treatment or control group. The treated participants were informed about our concept of the external costs of transport and presented with a graphic showing the external cost rates per mode (Figure 3). We explained that their allocated budget would be used to cover the external costs generated by their travel, and that any remaining funds in their account at the conclusion of the study was theirs to keep. To make sure everyone understood, we included two comprehension questions. Participants who answered these questions incorrectly saw the same information once more. Upon completing the intermediate survey, all information was emailed to the participants, including a link to explanatory documents.

The budget was personalized based on participants' individual average daily external costs during the observation period (as computed based on the simplified methodology explained above) plus an additional 20%. This buffer was included to reduce the likelihood that participants would exhaust their budget due to price-unrelated shocks during the observation or treatment period (mean reversion). Throughout the duration of the treatment period, participants in the treatment group received weekly e-mail summaries,

Figure 3: Marginal external costs by mode



Notes: The figure presents the marginal external costs by type and mode, as shown to the participants.

enabling them to track their external costs (and thus their payments). The costs were presented by mode of transport (Fig. 4b). The report also included an individual list of the costliest trips of the last week (Fig. 4c), the remaining budget, as well as the amount of valid tracking days (Fig. 4d). Person-days with tracking information of less than 10 hours in Switzerland (stays and travel together) were defined as missing; for these days, we deducted the personal average external costs of the observation period (see Fig. 4d) in order to reduce the potential of manipulation by simply switching off the app and to ensure that our analysis was based only on sufficiently tracked days. Section A.2 provides further robustness checks for the existence of strategic behavior related to the use of the tracking app.

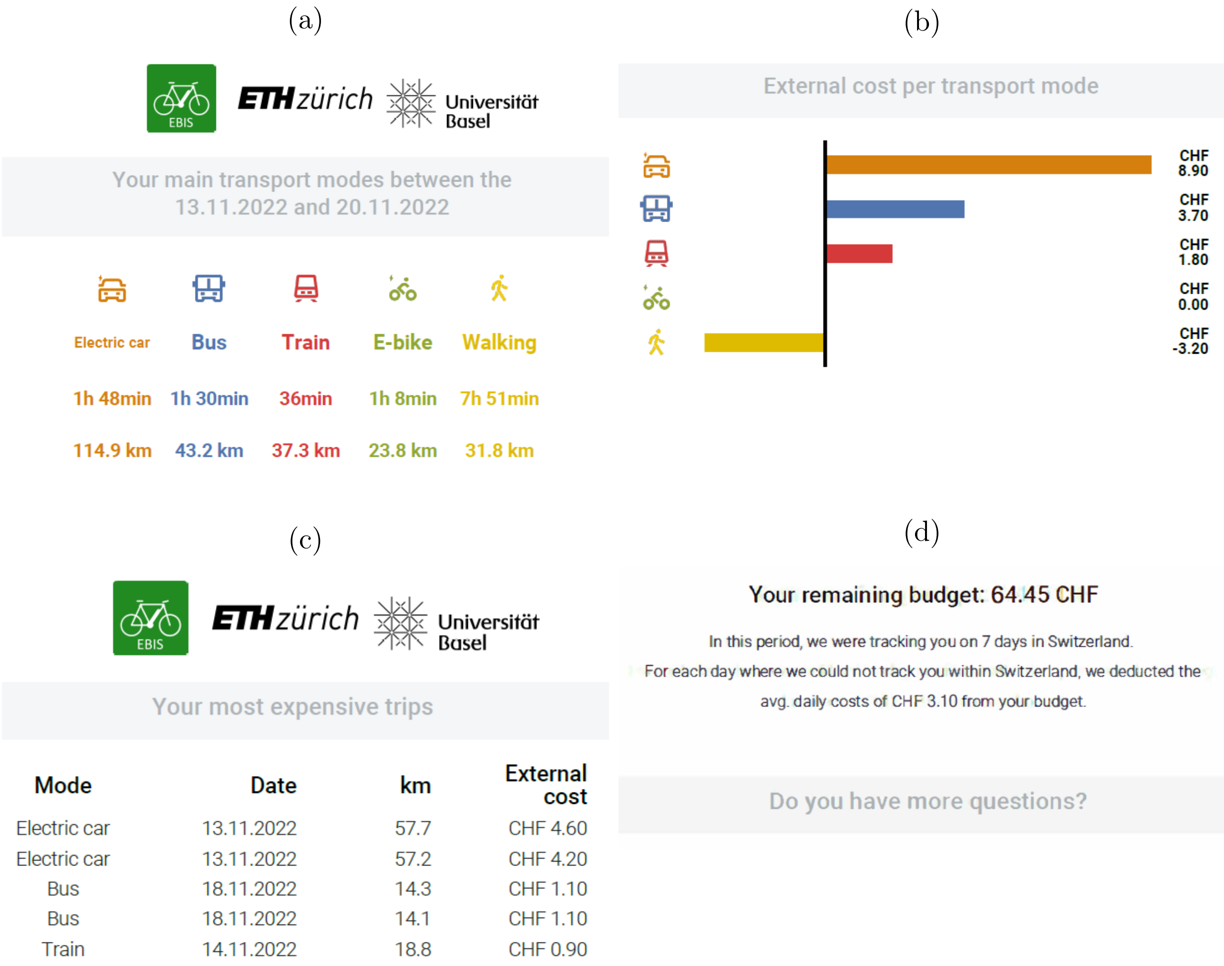
The control group continued to receive the information in Figure 4a in the form of weekly mobility reports during the entire experiment. In the intermediate survey, the control group was shown a graph summarizing the kilometer distances per mode of the baseline period (instead of information about external costs).

4 Empirical framework

This section describes our handling of the data and the empirical strategy for measuring the impact of the treatment.

Data preparation For the empirical analysis, we aggregate the data from the observed stage level to the person-day level. To increase the validity and accuracy of our results, we

Figure 4: Example of a weekly report



implemented a data cleaning process. It involved removing any implausible or obviously erroneous observations, which we believe to be primarily a result of measurement errors made by the app. We remove data points if the following conditions apply:

- Average daily speed exceeds 100 km/h for car, motorcycle, and public transport, 50 km/h for cycling, or 20 km/h for walking
- Total distance traveled exceeds 500 km/day for car, motorcycle, and public transport, 100 km/day for cycling, or 20 km/day for walking

Whenever any of these conditions was met, we removed the entire person-day from the analysis (as opposed to an individual trip). To limit potential bias from partially tracked days, we only use person-days with more than 10 hours of tracking within Switzerland (including stays).[13] We only included participants that delivered at least eight valid tracking days in the baseline period. Since some participants completed the intermediate survey with a delay, some people remained in the study for more than 70 days.

[13] On days that met this condition, we included all recorded travel, including trips outside Switzerland.

Despite the 20% buffer added, 15.6% of the individuals depleted their budget before the end of the study, most likely due to unforeseen changes in travel needs. In such cases, we informed participants that their travel budget would be increased once more (and one time only), scaled again according to their baseline travel needs. A total of 30 participants depleted their budget twice. The days with an exhausted travel budget are included in regressions but assigned a fixed effect to account for potential distortion from ineffective incentives.[14]

Regression analysis The randomized treatment yields an exogenous variation that can be directly used to identify causal treatment effects. The average treatment effect (ATE) is estimated by comparing means between treated and control observations using the following difference-in-differences (DiD) regression framework:

$$Y_{it} = c_0 + \tau \cdot DiD_{it} + \sum_{k=1}^{K} \beta_k \cdot x_{ik} \cdot DiD_{it} + \mu_i + \mu_t + \epsilon_{it} \tag{1}$$

The dependent variable is the outcome of interest for person $i \in (1, ..., N)$ on calendar day $t \in (1, ..., T)$. The main outcomes of interest are external costs (in CHF per day) and the distances by mode. The difference-in-differences term, DiD, is the product of a treatment group dummy (D_i) and a treatment period dummy (D_t). It equals one if the treatment is active for person i on a given day t, and zero otherwise. The ATE is given by the coefficient estimate $\hat{\tau}$. Due to rolling recruitment, the first day of the treatment falls on different calendar days for different people. To account for unobserved common shocks, we include individual-specific (μ_i) and date-specific (μ_t) fixed effects. The error term ϵ_{it} has an expected mean of zero and a variance of σ^2. We allow for correlation of the error within, but not across participants.[15]

To examine potential treatment effect heterogeneity (known as effect *moderation*), we include a k-dimensional vector of socio-demographic variables x and interact it with the treatment indicator. This allows us to examine, for instance, whether men respond more strongly to monetary incentives than women or whether income moderates the effect.

We are most interested in proportional effects to account for the fact that some people travel much more than others in absolute terms. This can be directly implemented by

[14] Omitting these 249 days from the regression would introduce systematic differential attrition across treatment groups.

[15] Our identification strategy touches on a recent literature that discusses the validity of the two-way fixed effects estimator in the context of "staggered" DiD designs (Callaway and Sant'Anna, 2021; de Chaisemartin and D'Haultfœuille, 2022; Sun and Abraham, 2021). However, in our setting, with rolling study participation and a large share of never-treated units, bad comparisons (i.e., late vs. early-treated) are of limited concern. Nevertheless, because the study spanned multiple seasons, used different recruitment channels and is subject to self-selection, it may be subject to dynamic treatment effects. To alleviate concerns in that regard, we confirm our results by applying the estimator suggested by Sun and Abraham (2021), which replicates our results almost exactly.

estimating Equation 1 using a Poisson Pseudo-Maximum Likelihood (PPML) model.[16] However, for the regressions that focus on total external costs, this approach would result in dropping all person-days that exhibit negative total external costs due to walking. To avoid this problem, we estimate these regressions in levels and then compute the proportional responses by dividing the coefficients (in CHF/d) by the potential average daily external costs of the treated group had they not been treated. This unobserved outcome is estimated by applying the same percentage change observed in the control averages to the treated pre-treatment average total external costs, thus following the spirit of the common trends assumption.

Thanks to the randomization of the treatment assignment, our setting is expected to satisfy the assumptions required for the causal identification of the ATE using the difference-in-differences estimator. Table 1 supports the comparability of the two study groups (and thus unconfounded assignment). Table A.1 and Table A.2 in the appendix provide an indication that the common trends hypothesis holds for distances and external costs during the observation period (and thus presumably also during the treatment period in the absence of treatment). We also find no indication for anticipation effects. Lastly, the identification of our treatment effects relies on the Stable Unit Treatment Value Assumption (SUTVA), which requires that there be no spillovers from the treated group to the control group. Since participants self-selected into the study, this assumption could be violated if individuals in the treatment and control groups know each other. To alleviate this issue, the treatment assignment process was adjusted to ensure that individuals sharing the same home address were assigned to the same group.

5 Results

5.1 Descriptive statistics

Table 4 provides an overview of the tracking data during the baseline period for the full sample. These numbers provide key statistics on the Swiss E-bike population and describe their travel behavior. The table aggregates data from Table 1 and presents averages at the participant-day level. Introducing a tax on the external costs of transport increases total marginal transport costs by about one-third. Of the tracked stages, 87% were confirmed by app users, while 6.5% were corrected. On average, we recorded 25 valid tracking days per person, covering 44 km and 92 minutes of travel per day.

The corresponding averages in the car-focused sample from Hintermann et al. (2024) are approximately 4.60 CHF/day in total external costs, 22.70 CHF/day in private costs,

[16] We prefer this specification over the log-linear model because it easily accommodates days with zero travel and mitigates potential bias from heteroskedasticity that can arise in the log-linear framework (see Santos Silva and Tenreyro, 2006).

Table 4: Summary of the tracking data in the baseline period

Variable	**Average**	**Std. Dev.**	**Unit**
Climate ext. costs	1.10	1.60	CHF/day
Congestion ext. costs	0.67	1.36	CHF/day
Health ext. benefits	-1.08	1.58	CHF/day
Health ext. costs	1.20	1.98	CHF/day
Accident ext. costs	1.46	1.74	CHF/day
Total ext. costs	**3.35**	**5.79**	**CHF/day**
Private costs	**10.20**	**16.40**	**CHF/day**
Car distance	27.06	47.44	km/day
Public transport distance	9.72	35.22	km/day
E-bike distance	4.95	10.64	km/day
Walking distance	1.91	2.53	km/day
Total distance	**43.95**	**56.49**	**km/day**
Total duration	**92.16**	**84.42**	**min/day**
Total stages	7.22	5.02	#/day
Confirmed	87.06	32.73	%
Corrected	6.53	15.78	%
Valid tracking days	25.07	4.98	days

Notes: Averages and standard deviations based on 27,196 recorded days during the baseline period.

and an average daily distance of 48 km. Despite these differences, likely due to higher car usage, the total time spent traveling per day is nearly identical (93 minutes).

5.2 Average treatment effects

Table 5 shows the ATE in absolute and relative terms. Column (1) shows the effect on total external costs. Introducing a Pigou-inspired tax reduces external costs by 0.213 CHF,[17] or 6.9% on average. The remaining columns report results per cost dimension and show that the largest effect consists in an increase in external health benefits. Accident-related externalities are not significantly reduced due to the substantial accident risk associated with E-biking. The relative reduction in total external costs implies an elasticity of 0.217 (s.e. 0.091) with respect to the average price increase of 31.6% due to the tax.[18]

Table 6 presents the ATEs on daily distances (measured in km), in total, and separately by mode. The price intervention has no significant effect on the total distance traveled. However, it leads participants to reduce their driving distance by 8.2% on average, while increasing the distance traveled by public transport, E-bike, and walking.

Figure 5 graphically illustrates the relative ATE by mode for both groups of outcome variables. It suggests that the overall decrease in total external costs is largely due to mode shift away from driving towards the other modes of transport.

[17] This coefficient was confirmed using the estimator by Sun and Abraham (2021), which results in a total effect of −0.207 CHF with a standard error of 0.088, and a p-value of 0.019, when grouping individuals according to the 106 unique dates when they entered the treatment phase.

[18] The corresponding standard errors are derived from a bootstrapping procedure with 1,000 draws.

Table 5: Average treatment effect on external costs

	(1) Total costs	(2) Climate	(3) Congestion	(4) Health benefits	(5) Health costs	(6) Accidents
ATE (CHF)	-0.213*	-0.049*	-0.068***	0.106***	-0.043	0.053
	(0.091)	(0.025)	(0.020)	(0.026)	(0.032)	(0.029)
adj. R^2	0.116	0.127	0.114	0.323	0.115	0.221
ATE (relative)	-0.069*	-0.050*	-0.116***	0.106***	-0.043	0.038
	(0.029)	(0.023)	(0.031)	(0.024)	(0.028)	(0.020)
Pseudo R^2	-	0.103	0.130	0.219	0.114	0.130
Clusters	1,085	1,085	1,085	1,085	1,085	1,085
N	61,410	61,410	61,410	61,410	61,410	61,410

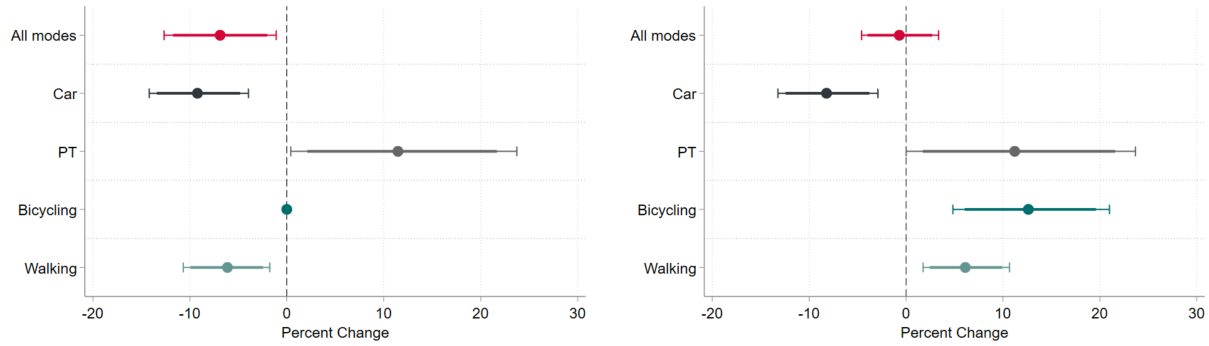
Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors (in parentheses) are clustered at the participant level. For total external costs, relative effects were calculated by dividing the ATE (in CHF) by the average of the control group during the treatment phase, using a bootstrap with 1,000 draws. For (2)-(6), relative effects are calculated using a PPML regression. All regressions include person and date fixed effects, as well as a dummy variable indicating days following the receipt of a negative travel budget in the mobility report.

Table 6: ATE on travel distance

	(1) Total distance	(2) Car	(3) Public transport	(4) E-Bike	(5) Walking
ATE (km)	-0.248	-1.935**	0.949	0.570**	0.118**
	(0.861)	(0.733)	(0.529)	(0.179)	(0.042)
adj. R^2	0.125	0.124	0.136	0.319	0.203
ATE (relative)	-0.007	-0.082**	0.112*	0.126**	0.061**
	(0.020)	(0.029)	(0.054)	(0.037)	(0.021)
Pseudo R^2	0.176	0.193	0.321	0.428	0.146
Clusters	1,085	1,085	1,085	1,085	1,085
N	61,410	61,410	61,410	61,410	61,410

Notes: * p<0.05, ** p<0.01, *** p<0.001. The dependent variable contains the distance traveled including zeroes aggregated to the person-day level. The ATE (km) coefficients show the ATE in kilometers. The relative coefficients were estimated using a PPML model. Standard errors (in parentheses) are clustered at the participant level. All regressions include person and date fixed effects, as well as a dummy variable indicating days following the receipt of a negative travel budget in the mobility report.

Figure 5: ATE on external costs and distances



(a) External costs (b) Travel distance

Notes: Graphical representation of the regression results in Table 5 and Table 6. The thick bars represent 90% confidence intervals, while the thin bars indicate 95% confidence intervals. Note that walking generates net benefits (i.e., negative external costs), while external costs associated with E-biking are set to zero.

5.3 Effect heterogeneity

So far, we have focused on the average treatment effect and the underlying mechanisms. Next, we examine effect moderation, i.e., how the effect varies with pre-treatment characteristics contained in the vector x in eq. (1) and over time.

Fast vs. regular E-bikes Table 7 indicates that the distance effect is primarily (though not exclusively) driven by S-pedelec owners. In contrast, the effect for regular E-bike owners (captured by the coefficient on *Treated*) is considerably smaller and statistically significant only for distances covered by public transport and on foot.

Table 7: Fast vs. slow E-bikes

	(1) Total distance	(2) Car	(3) Public transport	(4) E-Bike	(5) Walking
Treated	0.020	-0.049	0.169*	0.036	0.080**
	(0.027)	(0.036)	(0.071)	(0.053)	(0.027)
Treated x S-pedelec	-0.049	-0.067	-0.090	0.130*	-0.033
	(0.034)	(0.047)	(0.088)	(0.056)	(0.035)
Pseudo R^2	0.176	0.193	0.321	0.429	0.146
Clusters	1,085	1,085	1,085	1,085	1,085
N	61,410	61,410	61,410	61,410	61,410

Notes: * p<0.05, ** p<0.01, *** p<0.001. The coefficients were estimated using a PPML model, and the results show proportional effects. Standard errors (in parentheses) are clustered at the participant level. All regressions include person and date fixed effects, as well as a dummy variable indicating days following the receipt of a negative travel budget in the mobility report.

Pre-treatment transport behavior Table 8 presents the results from including interaction terms for pre-treatment distance shares of E-bikes and cars. The coefficient on *Treated* captures the treatment effect for those who had a prior E-bike or car share of zero. In the first part of Table 8, the significant and positive coefficients for these reference groups in columns (1)-(3) indicate that the relative increase in bicycle distance is strongest for individuals with an initially very low bicycle share, and decreases with the baseline share of this mode. Intuitively, people who already carry out most of their trips by bicycle cannot respond much to the pricing, whereas those that rarely use their E-bike can (and do) more easily increase usage in response to the treatment. Column (4) confirms that those with a lower pre-treatment E-bike share also reduce their car kilometers by more. Finally, columns (5) and (6) indicate that the shift away from driving is more pronounced for frequent drivers.

In the second part of Table 8, we estimate the effect in absolute terms (in kilometers) to address the concern that these results could be driven by similar absolute increases, which translate into much larger proportional changes for individuals with low pre-treatment mode shares. However, the absolute results confirm the pattern observed in the proportional ones: respondents with low baseline cycling shares exhibit significantly larger absolute increases than those with higher baseline shares. Similarly, columns (4)-(6) confirm that the reduction in absolute driving is larger for frequent drivers and for those that do not use their bicycle much during the pre-treatment period. Taken together, these results indicate that the mode shift is due to regular drivers driving less and infrequent cyclists cycling more.

Socio-demographic subgroups To examine effect heterogeneity with respect to socio-demographic characteristics, we engage in a multivariate analysis including several interaction terms. The interaction terms are chosen based on key socio-demographic variables identified as primary determinants of travel behavior in Hintermann et al. (2024). As shown in Table 9, the only significant interaction terms relate to E-bike distances. Column (4) indicates that individuals living in urban areas increase their bicycle distance significantly more, by 15.4%, compared to those in rural areas. Furthermore, individuals using faster E-bikes, such as S-pedelecs, increase their cycling distances by 17.7% more compared to those using standard (slower) E-bikes. We do not find evidence of effect heterogeneity with respect to the other modes.

We also conduct multivariate interactions for all types of external costs (see Table A.7 in the appendix). Most of the interaction terms are insignificant at conventional levels. The only exceptions are that people living in urban areas increase health benefits by 9% more (which is consistent with the result on E-bike distances) and that people above age 50 reduce accident-related external costs by 6.9% more than people below this age.

Table 8: ATE on travel distance with pre-treatment mode share interaction

	E-Bike distance (km/day)			Car distance (km/day)		
	(1)	(2)	(3)	(4)	(5)	(6)
Relative						
Treated	0.365***	0.002	0.414*	-0.125***	0.133	0.157
	(0.077)	(0.068)	(0.146)	(0.035)	(0.067)	(0.101)
Treated x E-bike share (pre)	-0.393**		-0.413**	0.314*		-0.041
	(0.162)		(0.204)	(0.137)		(0.179)
Treated x Car share (pre)		0.359	-0.054		-0.302***	-0.317**
		(0.171)	(0.214)		(0.097)	(0.126)
Pseudo R^2	0.429	0.429	0.429	0.193	0.193	0.193
Absolute (km)						
Treated	1.453***	-0.411	1.545***	-3.962***	4.586***	4.826**
	(0.203)	(0.331)	(0.444)	(0.960)	(1.189)	(1.788)
Treated x E-bike share (pre)	-3.972***		-4.072***	9.122***		-0.499
	(0.744)		(0.916)	(2.693)		(3.191)
Treated x Car share (pre)		2.006***	-0.142		-13.326***	-13.590***
		(0.518)	(0.608)		(2.213)	(2.636)
adj. R^2	0.320	0.319	0.320	0.124	0.124	0.124
Clusters	1,085	1,085	1,085	1,085	1,085	1,085
N	61,410	61,410	61,410	61,410	61,410	61,410

Notes: * p<0.05, ** p<0.01, *** p<0.001. The dependent variable contains the distance traveled including zeroes aggregated to the person-day level. The E-bike and car shares are pre-treatment average km-shares over all person-days. The relative coefficients were estimated using a PPML model, and the results show proportional effects. Standard errors (in parentheses) are clustered at the participant level. All regressions include person and date fixed effects, as well as a dummy variable indicating days following the receipt of a negative travel budget in the mobility report.

Table 9: Multivariate interactions: Distances

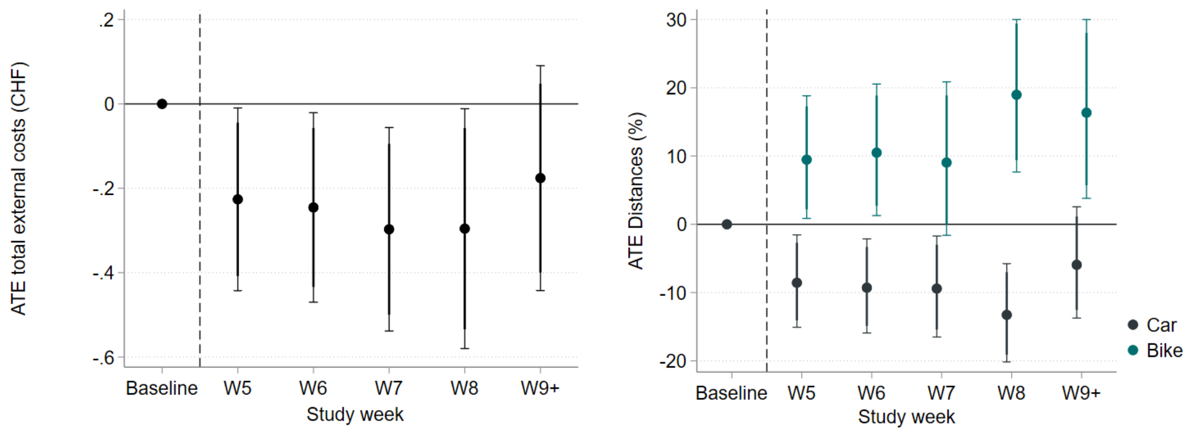
	(1) Total distance	(2) Car	(3) Public transport	(4) E-Bike	(5) Walking
Treated	0.081	0.025	0.521	0.029	0.094
	(0.060)	(0.078)	(0.234)	(0.139)	(0.065)
Treated x Male=1	-0.060	-0.021	-0.164	-0.062	0.004
	(0.033)	(0.046)	(0.092)	(0.064)	(0.034)
Treated x Age>=50	-0.047	-0.066	-0.014	-0.033	-0.029
	(0.035)	(0.050)	(0.094)	(0.065)	(0.036)
Treated x Tertiary educ.=1	0.006	0.012	0.009	-0.026	0.010
	(0.036)	(0.049)	(0.118)	(0.067)	(0.039)
Treated x HH size<3	-0.015	-0.015	-0.028	0.013	-0.032
	(0.035)	(0.050)	(0.092)	(0.068)	(0.037)
Treated x French=1	0.012	0.041	-0.074	-0.061	-0.037
	(0.048)	(0.067)	(0.132)	(0.085)	(0.049)
Treated x Urban=1	-0.007	-0.053	0.067	0.154*	-0.006
	(0.031)	(0.045)	(0.082)	(0.062)	(0.035)
Treated x PT reduction=1	0.012	-0.000	-0.169	-0.019	0.028
	(0.040)	(0.052)	(0.176)	(0.077)	(0.048)
Treated x S-pedelec=1	-0.036	-0.066	-0.069	0.177**	-0.042
	(0.035)	(0.049)	(0.095)	(0.060)	(0.037)
Pseudo R^2	0.176	0.193	0.321	0.429	0.146
Clusters	1,085	1,085	1,085	1,085	1,085
N	61,410	61,410	61,410	61,410	61,410

Notes: * p<0.05, ** p<0.01, *** p<0.001. The dependent variable contains the distance traveled restricted to positive observations aggregated to the person-day level. All dimensions include one omitted category. *Treated* is thus associated with an observation that has a zero for all included dummies. The coefficients were estimated using a PPML model, and the results show proportional effects. Standard errors (in parentheses) are clustered at the participant level. All regressions include person and date fixed effects, as well as a dummy variable indicating days following the receipt of a negative travel budget in the mobility report.

Variation over time Figure 6 presents DiD regression results with separate treatment dummies for each study week. This enables us to estimate separate treatment effects for each of the five weeks in the post-treatment phase, relative to the pre-treatment average. While there is some variation in the point estimates for both external costs and distances, these are not statistically different from one another, meaning we cannot reject the null hypothesis of an immediate and constant treatment effect.

The EBIS study spanned multiple seasons. Table 10 presents the treatment effects separately for individuals who started the RCT in autumn versus those who started in spring. The total effect in the autumn wave is captured by the coefficient on *Treated*, whereas the total effect for the spring wave is listed below. Table 10 shows that the observed effects are mostly driven by participants in the autumn wave, as none of the treatment effects in the spring wave are statistically significant, even though most point estimates have the expected sign (except for walking, which shows an effect close to zero). Given our data and recruitment strategy, we cannot determine whether the effect is genuinely absent in spring or whether participants recruited during this period, who were mostly from Zurich canton and primarily owners of fast E-bikes, are simply less price-responsive

Figure 6: Treatment effect dynamics



(a) External costs (CHF)

(b) Travel distance (%)

Notes: Both figures display results from a DiD-type regression in which the treatment dummy is replaced by an interaction between each post-treatment study week and the treated group. This approach is related to the event study design but differs in that all baseline period weeks are used as a control group to increase statistical power. All regressions include person and date fixed effects. The thick bars represent 90% confidence intervals, while the thin bars indicate 95% confidence intervals.

than earlier recruits. Additionally, since the spring treatment group consisted of only 157 individuals (compared to 420 in the autumn wave), the lack of significant effects may also be due to insufficient statistical power in the second wave.

5.4 Mechanisms

To identify the mechanisms that give rise to the treatment effect, we engage in a mediation analysis using the methodology developed by Baron and Kenny (1986), Kraemer et al. (2008), and Imai et al. (2010).

Given that a substantial part of the treatment effect seems to arise from a decrease in driving (see Figure 5), our candidate for the role of main mediator is car distance. We thus regress car distance on the treatment effect (to measure the effect of the treatment on driving) and include car distance as a control variable in a second regression in which we regress total external costs on the treatment indicator; we also include an interaction term between the treatment indicator and the mediator to account for the possibility that the relationship between the treatment and the outcome variable differs with the amount of driving.

Table 11 presents the estimates for the Average Direct Effect (ADE) as well as the Average Indirect Effect (AIE). The latter captures the effect via the assumed mediator, whereas the former measures the sum of all other mechanisms. The absence of a statistically significant ADE in column (1) suggests that driving explains the entire effect on external costs or, alternatively, that all other effects add up to zero. Column (2) shows that the

Table 10: Seasonality of the treatment effects

	(1) Total distance	(2) Car	(3) Public transport	(4) E-Bike	(5) Walking
Treated	-0.011	-0.103**	0.157*	0.146**	0.090***
	(0.025)	(0.034)	(0.064)	(0.043)	(0.025)
Treated x Spring	0.015	0.089	-0.129	-0.048	-0.088
	(0.044)	(0.064)	(0.121)	(0.079)	(0.049)
Treated +	0.004	-0.023	0.008	0.087	-0.007
Treated x Spring	(0.036)	(0.054)	(0.102)	(0.066)	(0.042)
Pseudo R^2	0.176	0.193	0.321	0.428	0.146
Clusters	1,085	1,085	1,085	1,085	1,085
N	61,410	61,410	61,410	61,410	61,410

Notes: * p<0.05, ** p<0.01, *** p<0.001. The dependent variable contains the distance traveled including zeroes aggregated to the person-day level. The coefficients were estimated using a PPML model, and the results show proportional effects. Standard errors (in parentheses) are clustered at the participant level. All regressions include person and date fixed effects, as well as a dummy variable indicating days following the receipt of a negative travel budget in the mobility report.

latter is the case: the external costs of transport are also affected by changes in public transport, bicycling, and walking, but these effects neutralize each other.

In column (3), we examine the mechanisms underlying the reduction in external congestion costs. Individuals essentially have two options to reduce these costs. They can (i) drive less or (ii) drive during less congested times. In theory, they could also drive in less congested areas, but we assume that the home and work locations are not changed due to a five-week treatment.

The significant and large AIE in column (3) indicates that the primary source of these reductions is indeed the decrease in car distances. However, the significant ADE suggests that individuals also shifted their car trips away from congested periods. In principle, both effects could take place at the same time if individuals replace car trips preferentially during congested times. However, even this interpretation is consistent with people not only adjusting the quantity of their driving but also the timing.

6 Discussion and conclusions

In this section, we discuss threats to internal and external validity, both of which are of first-order importance to interpret our results. We conclude with the policy implications of our work.

Internal validity Empirical studies often face challenges with validity due to non-random treatment assignment or the lack of a pure control group. Our field experiment

Table 11: Mediation analysis

	(1) Total ext. costs	(2) Total ext. costs	(3) Car congest. ext.
ADE	0.002	-0.009	-0.021*
	[-0.082,0.077]	[-0.083,0.061]	[-0.039,-0.001]
AIE (Car distance)	-0.215*	-0.218*	-0.047*
	[-0.371,-0.054]	[-0.376,-0.055]	[-0.081,-0.012]
AIE (Public transport distance)		0.033+	
		[-0.002,0.068]	
AIE (E-Bike distance)		-0.005**	
		[-0.009,-0.002]	
AIE (Walking distance)		-0.013**	
		[-0.023,-0.005]	
N	61,410	61,410	61,410

Notes: * p<0.05, ** p<0.01, *** p<0.001. The bounds show the 95%- percentile bootstrap confidence intervals (1,000 draws), which is the recommended choice for mediation analysis (Tibbe and Montoya, 2022). Columns (1) and (3) include car distance as the mediator variable. Column (2) incorporates all distances as separate mediators.

addresses both issues, enabling us to estimate causal treatment effects for our study sample. The assumptions required for the Difference-in-Differences approach have been demonstrated to hold to the extent testable, which supports the assumption that any difference during the treatment period is caused by the treatment itself.

Measurement error is of concern in this study due to the GPS-based data collection, which is imperfect, combined with the potential for individuals to modify the data. Appendix A.2 provides a detailed explanation of how alterations to the tracking data can be tested for. We find that treated individuals were not more likely to correct their imputed modes, but conditional on making a correction, they were more likely to correct their imputed modes away from car use, relative to the control group. This could be interpreted as evidence of cheating in the sense that (some) participants may have adjusted their reported modes strategically to reduce their external costs in the post-treatment phase. However, the absence of a measurable effect on corrections overall contradicts this interpretation, suggesting that participants did not make additional adjustments but simply paid more attention to car trips, as these were highlighted in the treatment e-mails. It remains therefore an open question whether the corrections were "honest", that is, whether an erroneously imputed car mode was accurately corrected to reflect the actual mode or strategically manipulated. Figure A.2 in the appendix shows that ignoring all participant corrections — whether truthful or otherwise — does not substantially weaken the results of the study. This also supports the reliability of the mode detection functionality in the tracking app. We conclude that whereas we cannot exclude the possibility that some participants manipulated the data in order to gain a financial advantage, these manipulations were not important enough to drive our results.

External validity The validity of field experiment results across contexts can be assessed through four points, as outlined by List (2020): selection into the study or into treatment groups, differential attrition and observability, the naturalness of the experimental setting, and the scalability of the findings. In the following, we discuss each of these in turn.

Selection While our experimental framework eliminates selection into treatment, potential bias may arise from the targeted recruitment process and participants' willingness to engage in the smartphone tracking study. Table 1 presents the socio-demographic characteristics of the final RCT sample in comparison to the relevant subpopulation from the Swiss Microcensus on transport (2023). Key differences include a lower proportion of individuals aged 18-30, and a higher proportion of males, urban residents, and fast E-bike owners in the RCT sample. Given that these characteristics modify the treatment effect (see section 5.3), it is to be expected that our results do not directly apply to all E-bikers in Switzerland. Knowledge of the conditional treatment effects allows us to compute the expected effect for a different sample, provided that there is common support among the key determinants. For example, we have computed the differential response of S-pedelec owners relative to owners of regular E-bikes, and it is therefore straightforward to compute the ATE for a sample in which the shares of these E-bikes are different. The same is true for all observable characteristics that we find to be modifiers of the effect.

Assessing a potential bias due to the self-selection into the smartphone tracking study is difficult and would require a variation in the incentive payment, which was not done here. But the participants are likely to be systematically different to those who declined participation. To the extent that the unobservable characteristics that co-determine participation are modifiers of the treatment effect, our results will be biased. Table 1 shows the sample characteristics of all individuals who were considered eligible during the introduction survey. Especially individuals aged 31-50 or 66-87 years, with secondary level education, or living with another person in their household were more reluctant to participate in the tracking study after completing the first survey. We acknowledge this limitation of our study and cannot determine the magnitude or direction of this self-selection bias. On the other hand, if some form of transport pricing were to be implemented in a voluntary manner in exchange for some other tax relief, our sample would arguably constitute a very good basis to predict the treatment effect of such a program.

Attrition and observability Attrition is hardly avoidable in a field experiment, and it poses a concern when participants who complete the study differ from those who do not. If completion probability is related to key outcome variables, such as car distance, treatment effects may be biased. Another concern is that attrition varies by treatment status. Both potential sources of bias can be assessed through a regression analyzing the determinants of attrition. Appendix A.3 demonstrates that neither treatment status nor outcome variables significantly influence the observability of individuals in the treatment phase, implying that our results are not driven by differential attrition.

Naturalness A key consideration when interpreting experimental results is the naturalness of the task under consideration. In this study, we observed people in their regular environment as they make everyday transport choices. This is a key advantage of a field experiment, relative to more artificial settings such as laboratory experiments.

The most "unnatural" part of the experiment is the implementation of the Pigovian tax as a deduction from a travel budget that we previously assigned to the participants. Although equivalent in strictly microeconomic terms, it is unclear whether the (psychological) effect of reducing this budget is equivalent to that of imposing an actual tax. Without governmental authority, the approach chosen in this study is arguably the most practical approximation of an actual tax. Since participants were not obliged to pay for external costs that exceeded their budget, our study can only estimate a substitution effect, as the income effect of the tax is compensated for by the individualized budget. Furthermore, there may be a behavioral distinction between receiving less money and paying taxes directly from one's own assets. Thaler and Johnson (1990) suggest that people often treat gambling money differently from their regular income, blending prior gains with subsequent losses and viewing losses smaller than the initial gain as less significant. This can encourage risk-seeking behavior, potentially leading to an underestimation of the effect that would result from transport pricing which is deducted from households' actual income.

Scaling An important aspect is the extent to which the experimental results are likely to scale to different populations. Scaling can take three qualitatively different forms. First, horizontal scaling determines how the treatment effects generalize to different samples in the sense that the same experiment would be carried out elsewhere. Regional variations, such as the quality of the Swiss public transport network, may limit horizontal scalability in our context. On the other hand, our main effect consists of a reduction in driving combined with an increase in E-biking, which could also take place in the absence of public transport. The quality of cycling infrastructure and overall cycling safety is another matter. For this reason, we do not expect our results to translate to settings in which bicycling is perceived as much less safe than in Switzerland. Although the study covers a broad sample of adults aged 18 and older, it focuses on E-bike users who also regularly use a car. The effect will likely be smaller for people who own neither a car nor an E-bike, as the main substitution pathway identified in our study is not available. Seasonal variation and climatic conditions are another reason for why our results may not be directly applicable everywhere. Our treatment effect is a weighted average of the effects observed in autumn and spring. Although no significant differences were found, the difference in point estimates suggests a dependency on the weather and climate.

Second, vertical scaling considers the impact of extending the treatment to a larger share of the population. Implementing a nationwide Pigovian tax on transport externalities would introduce general equilibrium effects. The policy reduces car kilometers traveled and encourages a shift to more sustainable modes. This would presumably lead to less congested roads which could, in turn, attract additional drivers and discourage some individuals from cycling due to increased bicycle lane traffic. These feedback loops could shift the equilibrium, affecting external costs and thus the marginal rates of external costs that determine the tax itself. Such dynamic adjustments cannot be captured within the scope of this field experiment and would have to be modeled separately.

Third, the results should also be scalable with respect to time. The five-week tax exposure in this study likely captures short-term behavioral changes, focusing on intensive-margin effects, but may not account for long-term decisions such as changes in mode tool ownership or residential location. Consequently, our estimates might understate the effects of prolonged exposure to a Pigovian tax on transport externalities. In absence of direct evidence, long-term effects may be approximated from studies on fuel price elasticity, where findings by Goodwin et al. (2004) indicate that long-run elasticities exceed short-run elasticities by a factor of 2 to 3.

Policy implications This study shows the effectiveness of reducing the external costs of transport through the implementation of a first-best Pigovian tax. While overall travel distances remain unchanged, we observe a substantial shift in travel distances away from cars and towards public transport and E-bikes, and especially towards S-pedelecs. We also find indirect evidence of drivers shifting travel away from congested time windows. As E-bikes likely continue to grow in popularity, our findings highlight their potential to play a significant role in shaping future transport policy.

In a future update of this paper we will include the welfare implications of such a policy, considering not only the benefits of the reduced external costs of transport but also the utility that individuals experience when they change their transport behavior in response to the pricing.

References

Agarwal, S. and Koo, K. M. (2016). Impact of electronic road pricing (erp) changes on transport modal choice. *Regional Science and Urban Economics*, 60:1–11.

Andersson, A. S., Adell, E., and Hiselius, L. W. (2021). What is the substitution effect of e-bikes? a randomised controlled trial. *Transportation research part D: transport and environment*, 90:102648.

Baron, R. M. and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, 51(6):1173.

Ben-Elia, E. and Ettema, D. (2011). Rewarding rush-hour avoidance: A study of commuters' travel behavior. *Transportation Research Part A: Policy and Practice*, 45(7):567–582.

Bieler, C., Sutter, D., Lieb, C., Amacher, M., and Sommer, H. (2019). Externe effekte des verkehrs 2015 aktualisierung der berechnungen von umwelt-, unfall- und gesundheitseffekten des strassen-, schienen-, luft- und schiffsverkehrs 2010 bis 2015. Technical report, Bundesamt für Raumentwicklung ARE.

Bigazzi, A. and Wong, K. (2020). Electric bicycle mode substitution for driving, public transit, conventional cycling, and walking. *Transportation Research Part D: Transport and Environment*, 85:102412.

Böcker, L., Dijst, M., and Faber, J. (2016). Weather, transport mode choices and emotional travel experiences. *Transportation Research Part A: Policy and Practice*, 94:360–373.

Börjesson, M. and Kristoffersson, I. (2018). The swedish congestion charges: Ten years on. *Transportation Research Part A: Policy and Practice*, 107:35–51.

Bourne, J. E., Cooper, A. R., Kelly, P., Kinnear, F. J., England, C., Leary, S., and Page, A. (2020). The impact of e-cycling on travel behaviour: a scoping review. *Journal of transport & health*, 19:100910.

Buehler, R. and Pucher, J. (2021). *Cycling for Sustainable Cities*. The MIT Press.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

Castro, A., Kahlmeier, S., and Gotschi, T. (2018). Exposure-adjusted road fatality rates for cycling and walking in european countries. Technical report, OECD. `www.itf-oecd.org/exposure-adjusted-road-fatality-rates-cycling-and-walking-european-countries` (last accessed: 23-08-2024).

Charging, I. (2019). Sustainable transport infrastructure charging and internalisation of transport externalities: main findings.

Council, P. S. R. (2008). Traffic choices study. Technical report, `https://www.psrc.org/sites/default/files/2022-03/traffic-choices.pdf` (last accessed: 30-12-2024).

de Chaisemartin, C. and D'Haultfœuille, X. (2022). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: a survey. *The Econometrics Journal*, 26(3):C1–C30.

De Kruijf, J., Ettema, D., Kamphuis, C. B., and Dijst, M. (2018). Evaluation of an incentive program to stimulate the shift from car commuting to e-cycling in the netherlands. *Journal of Transport & Health*, 10:74–83.

Ecoplan-INFRAS (2024). Externe effekte des verkehrs 2021. umwelt-, unfall- und gesundheitseffekte des strassen-, schienen-, luft- und schiffsverkehrs. Technical report, Bundesamt für Raumentwicklung. https://www.are.admin.ch/are/de/home/mobilitaet/grundlagen-und-daten/kosten-und-nutzen-des-verkehrs.html.

Eliasson, J. (2009). A cost–benefit analysis of the stockholm congestion charging system. *Transportation Research Part A: Policy and Practice*, 43(4):468–480.

Eliasson, J. (2021). Efficient transport pricing–why, what, and when? *Communications in Transportation Research*, 1:100006.

Eurostat (2021). Oecd regional development studies applying the degree of urbanisation a methodological manual to define cities, towns and rural areas for international comparisons. Technical report, European Commission, the Food and Agriculture Organization of the United Nations (FAO), the United Nations Human Settlements Programme (UN-Habitat), the International Labour Organization (ILO), the Organisation for Economic Co-operation and Development (OECD) and The World Bank.

Fishman, E. and Cherry, C. (2016). E-bikes in the mainstream: Reviewing a decade of research. *Transport reviews*, 36(1):72–91.

Ghanem, D., Hirshleifer, S., and Ortiz-Becerra, K. (2023). Testing attrition bias in field experiments. *Journal of Human resources*.

Gibson, M. and Carnovale, M. (2015). The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics*, 89:62–73.

Goodwin, P., Dargay, J., and Hanly, M. (2004). Elasticities of road traffic and fuel consumption with respect to price and income: a review. *Trans. reviews*, 24(3):275–292.

Greene-Roesel, R., Castiglione, J., Guiriba, C., and Bradley, M. (2018). Bart perks: using incentives to manage transit demand. *Transportation Research Record*, 2672(8):557–565.

Hahn, R. W., Metcalfe, R. D., and Tam, E. (2024). Welfare estimates of shifting peak travel. *National Bureau of Economic Research*. Working paper.

Heinen, E. and Handy, S. (2021). Programs and Policies Promoting Cycling. In *Cycling for Sustainable Cities*. The MIT Press.

Heinen, E., Van Wee, B., and Maat, K. (2010). Commuting by bicycle: an overview of the literature. *Transport reviews*, 30(1):59–96.

Heinonen, S., de Freitas, L. M., Meister, A., Schwab, L., Roth, J., Hintermann, B., Götschi, T., and Axhausen, K. (2024). The e-biking in switzerland (ebis) study: Methods and dataset. *Transportation*, pages 1–24.

Hintermann, B., Schoeman, B., Molloy, J., Götschi, T., Castro, A., Tchervenkov, C., Tomic, U., and Axhausen, K. W. (2024). Pigovian transport pricing in practice. *WWZ Discussion Papers*, 2021/11 (update).

Imai, K., Keele, L., and Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, 25:51–71.

Isaksen, E. T. and Johansen, B. G. (2021). Congestion pricing, air pollution, and individual-level behavioral responses. *Available at SSRN 3832230*.

Karlström, A. and Franklin, J. P. (2009). Behavioral adjustments and equity effects of congestion pricing: Analysis of morning commutes during the stockholm trial. *Transportation Research Part A: Policy and Practice*, 43(3):283–296.

Kraemer, H. C., Kiernan, M., Essex, M., and Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the baron & kenny and macarthur approaches. *Health Psychology*, 27(2S):S101.

Krebs, P. and Balmer, U. (2015). Fair and efficient: The distance-related heavy vehicle fee (hvf) in switzerland. Technical report, Federal Office for Spatial Development (ARE). `https://www.are.admin.ch/are/en/home/mobility/data/distance-related-heavy-vehicle-fee--hvf-.html` (last accessed: 08-21-2024).

Kreindler, G. (2024). Peak-hour road congestion pricing: Experimental evidence and equilibrium implications. *Econometrica*, 92(4):1233–1268.

Kroesen, M. (2017). To what extent do e-bikes substitute travel by other modes? evidence from the netherlands. *Transportation Research Part D: Transport and Environment*, 53:377–387.

Leape, J. (2006). The london congestion charge. *Journal of economic perspectives*, 20(4):157–176.

List, J. A. (2020). Non est disputandum de generalizability? a glimpse into the external validity trial. Technical report, National Bureau of Economic Research.

Martin, L. and Thornton, S. (2018). To drive or not to drive? a field experiment in road pricing. Working paper.

McQueen, M., MacArthur, J., and Cherry, C. (2020). The e-bike potential: Estimating regional e-bike impacts on greenhouse gas emissions. *Transportation Research Part D: Transport and Environment*, 87:102482.

Metropia (2024). Incentivizing active & shared travel pilot program: Final report. Technical report, Metropolitan Transportation Commission.

Molloy, J., Castro Fernández, A., Götschi, T., Schoeman, B., Tchervenkov, C., Tomic, U., Hintermann, B., and Axhausen, K. W. (2020). A national-scale mobility pricing experiment using gps tracking and online surveys in switzerland: Response rates and survey method results. *Arbeitsberichte Verkehrs-und Raumplanung*, 1555.

Molloy, J., Tchervenkov, C., and Axhausen, K. W. (2021). Estimating the external costs of travel on gps tracks. *Transportation Research Part D: Transport and Environment*, 95:102842.

Moser, C., Blumer, Y., and Hille, S. L. (2018). E-bike trials' potential to promote sustained changes in car owners mobility habits. *Environmental Research Letters*, 13:044025.

Neves, A. and Brand, C. (2019). Assessing the potential for carbon emissions savings from replacing short car trips with walking and cycling using a mixed gps-travel diary approach. *Transportation Research Part A: Policy and Practice*, 123:130–146.

Nielsen, O. A. (2004). Behavioral responses to road pricing schemes: Description of the danish akta experiment. *Intelligent Transportation Systems*, 8(4):233–251.

Nilsson, P., Tarduno, M., and Tebbe, S. (2024). Road pricing with green vehicle exemptions: Theory and evidence.

Olszewski, P. and Xie, L. (2005). Modelling the effects of road pricing on traffic in singapore. *Transportation Research Part A: Policy and Practice*, 39(7-9):755–772.

Philips, I., Anable, J., and Chatterton, T. (2022). E-bikes and their capability to reduce car co2 emissions. *Transport Policy*, 116:11–23.

Pigou, A. C. (1920). *The economics of welfare*. Macmillan & Co.

Plazier, P. A., Weitkamp, G., and van den Berg, A. E. (2017). "cycling was never so easy!" an analysis of e-bike commuters' motives, travel behaviour and experiences using gps-tracking and interviews. *Journal of transport geography*, 65:25–34.

Reck, D. J., Martin, H., and Axhausen, K. W. (2022). Mode choice, substitution patterns and environmental impacts of shared and personal micro-mobility. *Transportation Research Part D*, 102:103134.

Rosenfield, A., Attanucci, J. P., and Zhao, J. (2020). A randomized controlled trial in travel demand management. *Transportation*, 47(4):1907–1932.

Sacchi, R. and Bauer, C. (2023). mobitool v3.0. Technical report, Paul Scherrer Institut. `www.mobitool.ch/de/tools/mobitool-faktoren-v3-0-25.html`.

Santos, G. (2017). Road fuel taxes in europe: Do they internalize road transport externalities? *Transport Policy*, 53:120–134.

Santos Silva, J. M. C. and Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Simeonova, E., Currie, J., Nilsson, P., and Walker, R. (2021). Congestion pricing, air pollution, and children's health. *Journal of Human Resources*, 56(4):971–996.

Small, K. and Verhoef, E. T. (2007). *The economics of urban transportation*. Routledge.

Smith, M., Hosking, J., Woodward, A., Witten, K., MacMillan, A., Field, A., Baas, P., and Mackie, H. (2017). Systematic literature review of built environment effects on physical activity and active transport–an update and new findings on health equity. *International journal of behavioral nutrition and physical activity*, 14:1–27.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics*, 225(2):175–199.

Sun, Q., Feng, T., Kemperman, A., and Spahn, A. (2020). Modal shift implications of e-bike use in the netherlands: Moving towards sustainability? *Transportation Research Part D: Transport and Environment*, 78:102202.

Swiss Federal Office of Spatial Development (2023). Mobility and transport microcensus 2021. Technical report, Swiss Federal Statistical Office. `https://www.are.admin.ch/are/en/home/mobility/data/mtmc.html`.

Thaler, R. H. and Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science*, 36(6):643–660.

Tibbe, T. D. and Montoya, A. K. (2022). Correcting the bias correction for the bootstrap confidence interval in mediation analysis. *Frontiers in psychology*, 13:810258.

Transport for London (2007). Central london congestion charging: Impacts monitoring: Fifth annual report, july 2007. Technical report, Transport for London.

Tsirimpa, A., Polydoropoulou, A., Pagoni, I., and Tsouros, I. (2019). A reward-based instrument for promoting multimodality. *Transportation research part F: traffic psychology and behaviour*, 65:121–140.

Van Essen, H., Van Wijngaarden, L., Schroten, A., Sutter, D., Bieler, C., Maffii, S., Brambilla, M., Fiorello, D., Fermi, F., Parolin, R., et al. (2019). *Handbook on the external costs of transport, version 2019*. Number 18.4 K83. 131. CE Delft.

Verhoef, E. T. (2000). The implementation of marginal external cost pricing in road transport. *Papers in Regional Science*, 79(3):307–332.

Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review*, pages 251–260.

Yang, J., Purevjav, A.-O., and Li, S. (2020). The marginal cost of traffic congestion and road pricing: evidence from a natural experiment in beijing. *American Economic Journal: Economic Policy*, 12(1):418–53.

Yang, L., Sahlqvist, S., McMinn, A., Griffin, S. J., and Ogilvie, D. (2010). Interventions to promote cycling: systematic review. *Bmj*, 341.

Appendices

A Additional tables and figures

A.1 Identification strategy

A canonical test for the common trends assumption involves comparing the linear trends of the control and treatment groups prior to treatment. The regressions presented below include only days within the observation period. The absence of a significant "Treated x Study day" interaction term supports the assumption of parallel (linear) trends between the two groups.

Table A.1: Common pre-treatment linear trends in distances

	(1) Total distance	(2) Car	(3) Motorcycle	(4) Public transport	(5) Bicycle	(6) Walking
Treated	-0.497	-1.100	-0.045	0.671	-0.163	0.141
	(1.851)	(1.540)	(0.245)	(1.152)	(0.428)	(0.097)
Study day	-0.012	-0.065	-0.019	0.026	0.049**	-0.003
	(0.082)	(0.076)	(0.012)	(0.043)	(0.019)	(0.004)
Treated × Study day	-0.044	-0.052	0.004	-0.001	0.007	-0.001
	(0.078)	(0.071)	(0.010)	(0.045)	(0.016)	(0.004)
adj. R^2	0.005	0.013	-0.001	0.001	0.075	0.009
Clusters	1,084	1,084	1,084	1,084	1084	1,084
N	27,147	27,147	27,147	27,147	27,147	27,147

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors (in parentheses) are clustered at the participant level. All regressions are based solely on pre-treatment data and include date fixed effects.

Table A.2: Common pre-treatment linear trends in external costs

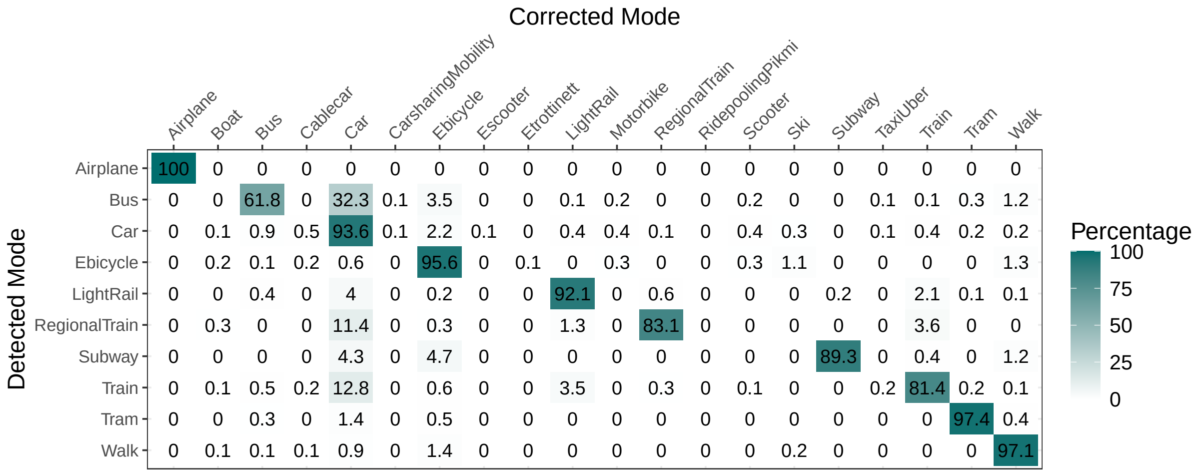
	(1) Total Ext.	(2) Environm. Ext.	(3) Cong. Ext.	(4) Health Ext.	(5) Accid. Ext.
Treated	-0.122	-0.038	-0.011	-0.035	-0.038
	(0.197)	(0.053)	(0.043)	(0.102)	(0.068)
Study day	-0.008	-0.000	0.001	-0.011*	0.002
	(0.010)	(0.003)	(0.002)	(0.005)	(0.003)
Treated × Study day	-0.004	-0.001	-0.001	-0.002	0.001
	(0.009)	(0.002)	(0.002)	(0.004)	(0.003)
adj. R^2	0.010	0.008	0.009	0.037	0.035
Clusters	1,084	1,084	1,084	1,084	1,084
N	27,147	27,147	27,147	27,147	27,147

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors (in parentheses) are clustered at the participant level. All regressions are based solely on pre-treatment data and include date fixed effects.

A.2 Tracking accuracy and strategic corrections

Tracking accuracy Participants had the option to correct the automatically detected travel mode. Figure A.1 displays the confusion matrix comparing the automatically detected modes with the corrections, for confirmed trip stages from non-treated individuals only. The tracking app did not automatically detect certain modes such as boats or cable-cars, resulting in a higher number of corrected modes.

Figure A.1: Confusion matrix of mode detection



Notes: Confusion matrix for mode detection among confirmed trip stages. The values represent percentage shares relative to the total number of stages classified as respective mode.

The confusion matrix demonstrates exceptionally high hit rates for most modes, with values generally exceeding 90%. However, the mode detection algorithm exhibits notable challenges in distinguishing bus trips from car trips, with 32.3% of bus trips being misclassified as car trips. A similar issue arises in the detection of (regional) train trips, albeit to a much lesser extent.

Strategic corrections The results of our study reaffirm that individuals respond to financial incentives. This raises the possibility that some treated participants may have adjusted their tracking diaries to maximize their financial rewards from the study. Such strategic modifications, diverging from genuine corrections as illustrated in Figure A.1, pose a threat to the internal validity of our findings and must be carefully examined. Our study design presents two potential avenues for participants to manipulate their behavior to maximize financial rewards. First, by falsely altering the detected travel mode to one associated with lower external costs. Second, by deactivating GPS tracking during trips with higher costs. Notably, the tracking app does not allow users to modify recorded GPS points or manually add trips. This design effectively minimizes the risk of manipulation, as only directly observed trips are included in the analysis. Both concerns can be addressed by analyzing the enriched GPS tracks. To assess the likelihood of untruthful corrections, we conduct a series of DiD regressions, replicating the main ATE regressions, using the percentage of corrected tracks per person-day as the outcome variable. These outcomes are expected to remain relatively stable before and after treatment, as they are not influenced by changes such as a reduction in car trips resulting from the treatment. Cheating behavior among treated individuals would be indicated by a significantly higher correction percentage during the treatment period. As shown by column (1) in Table A.3, no such pattern is observed. However, the mode-specific regressions in columns (3) and (4) reveal a significantly higher correction percentage away from car trips and, to a lesser extent, towards walking. Column (3) shows that participants in the treatment group correct 3.5 percentage points more car trips during the treatment period compared to the control group (11.9% vs. 8.4%). Column (5) shows no significant difference in the percentage of tracks marked as "completely misdetected" by participants. To investigate strategic fiddling through GPS deactivation, we examine the geographic distance between the final GPS point of one stage and the initial point of the subsequent stage, even if interrupted by a stay. Had treated individuals attempted to reduce external costs using this strategy, a significant difference in these distances should be observed. However, column (6) reveals no such evidence.

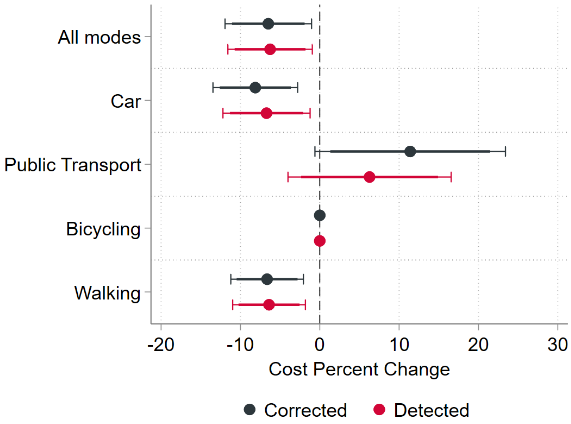
Table A.3: Regressions to detect cheating behavior

	Corrections					
	Overall (%) (1)	To car (%) (2)	From car (%) (3)	To walking (%) (4)	Deletions (%) (5)	Spatial jumps (km) (6)
Treated	0.296	-0.261	3.498***	0.166*	0.090	1.381
	(0.280)	(0.384)	(0.635)	(0.078)	(0.175)	(1.037)
adj. R^2	0.172	0.207	0.201	0.131	0.220	0.019
Clusters	1,085	1,085	1,085	1,085	1,085	1,085
N	57,814	36,867	38,969	50,735	57,814	61,410

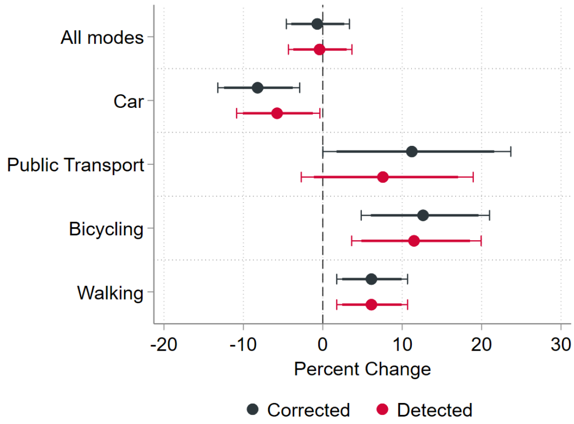
Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors (in parentheses) are clustered at the participant level. For columns (1)–(5), the outcome variable is defined as the percentage (%) of corrected tracks among all tracks of the respective mode on a given day. Cheating behavior among treated individuals might be indicated by a significantly higher correction percentage during the treatment period. The coefficients represent differences in percentage points. Column (5) examines the percentage of tracks marked as "completely misdetected". The outcome in column (6) is the total daily distance of spatial gaps (as the crow flies) in the GPS data.

The observed increase in corrections away from cars may suggest some degree of strategic behavior among treated individuals. However, it could just as easily reflect that treated participants put in more effort due to the larger financial incentive, leading to a higher correction rate. Fortunately, the comprehensive tracking data enables us to analyze raw, uncorrected data alongside corrected data. Using raw tracking data, based solely on automatically detected modes, eliminates any potential strategic reporting but also disregards genuine corrections by participants. The true effect of the experiment likely lies between these two data versions. Figure A.2 compares the results of our main ATE regressions using corrected and raw data. The comparison shows only minor differences, with no impact on the 5% significance level, except for the ATE on public transport distance.

Figure A.2: Comparison of ATEs with and without corrections



(a) External costs



(b) Distance

Notes: Comparison of the treatment effects accounting for participant corrections (black) versus the effects when all corrections are disregarded (red). The thick bars represent 90% confidence intervals, while the thin bars indicate 95% confidence intervals.

A.3 Differential attrition and observability

Differential attrition, where participants systematically drop out of a study based on treatment status or outcome variables, poses a key concern in field experiments (Ghanem et al., 2023). To address this concern, we conduct a determinants-of-attrition test to verify that attrition in our study is not systematically related to treatment assignment or key outcomes. To assess attrition, we define a variable that measures the number of days an individual was observable (i.e., provided a valid tracking day) during each study period. This measure is normalized by dividing it by the total possible observable days for each individual: at least 28 days for the baseline and 35 days for the treatment period. We estimate regressions for two outcome variables, "Observable (days)" and "Observable (%)", with results presented in Table A.4. The explanatory variables include study group assignment, as well as the individual's average baseline values for the main outcome variables, "external costs" and "car distance".

Table A.4: Determinants-of-attrition test

	External costs			Car distance		
	Observable (days)	Observable (%)		Observable (days)	Observable (%)	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.242	0.007	0.007	0.231	0.007	0.007
	(0.350)	(0.010)	(0.010)	(0.350)	(0.010)	(0.010)
Avg. baseline external costs	−0.144	−0.004	−0.003			
	(0.077)	(0.002)	(0.002)			
Baseline observability (%)			0.319***			0.318***
			(0.031)			(0.031)
Avg. baseline car distance				−0.019*	−0.001*	−0.0003
				(0.009)	(0.0003)	(0.0003)
Constant	31.889***	0.911***	0.643***	31.919***	0.912***	0.644***
	(0.370)	(0.011)	(0.028)	(0.366)	(0.010)	(0.028)
Adjusted R^2	0.002	0.002	0.092	0.002	0.002	0.092
N	1,085	1,085	1,085	1,085	1,085	1,085

Notes: * p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses. The table presents regressions for two outcome variables, "Observable (days)" and "Observable (%)". Observability is defined as the number of days an individual provided valid tracking data during each study period, normalized by the total possible observable days: at least 28 days for the baseline period and 35 days for the treatment period. Explanatory variables include study group assignment and the individual's average baseline values for the main outcome variables, "external costs" and "car distance".

The table reveals no significant differences in observability across treatment groups in any of the regressions. Furthermore, we do not find any relationship between baseline total external costs and the observability measures. However, a correlation does emerge between average baseline car distances and observability in the treatment period, depicted in columns (4) and (5). This relationship vanishes once we control for baseline observability as in column (6). Based on these findings, we conclude that there is no evidence of differential attrition in our study, reinforcing the internal validity of our main results.

A.4 Weather controls

Weather conditions significantly influence mode choice, particularly for active transport modes (Böcker et al., 2016). However, in our DiD framework, such factors should not bias the results. The random assignment to treatment or control groups ensures that weather effects are balanced across groups, thereby minimizing their influence on the estimated treatment effects. This assumption holds more robustly with larger sample sizes. Given our finite sample size, we augment the tracking data with high-resolution weather information from MeteoSwiss, including temperature and precipitation (in mm/h) at a 1 x 1 km spatial resolution. Following the methodology of Hintermann et al. (2024), we include temperature in two forms to account for the distinct effects of unusually hot and cool days:

$$\text{Heat}_{jt} = \max\{t_{jt}^{max} - 25,\, 0\}$$
$$\text{Cold}_{jt} = \max\{7 - t_{jt}^{min},\, 0\}$$

where t_{jt}^{max} and t_{jt}^{min} denote the daily maximum and minimum temperatures, respectively, at the location where a trip begins. For each trip j, these variables capture the positive deviations of daily temperature above 25 (heat) and below 7 (cold) degrees Celsius. Heat, cold and precipitation averages are then computed across all trips made by individual i on date t. On valid tracking days with no recorded trips, we use weather conditions at the last recorded location to impute values. Table A.5 reports the main regression results for distances traveled, incorporating these weather controls.

Table A.5: ATE of distances controlling for weather

	(1) Total distance	(2) Car	(3) Public transport	(4) E-Bike	(5) Walking
Treated	-0.219	-1.926**	0.961	0.572**	0.119**
	(0.861)	(0.734)	(0.529)	(0.179)	(0.041)
Heat day	-4.554**	-1.258	-2.267***	-0.263	-0.148*
	(1.580)	(1.257)	(0.687)	(0.336)	(0.066)
Cold day	-0.099	-0.011	-0.095*	0.004	-0.001
	(0.085)	(0.071)	(0.046)	(0.012)	(0.004)
Precipitation	-0.061	0.002	-0.004	-0.043***	-0.015***
	(0.088)	(0.076)	(0.049)	(0.012)	(0.004)
adj. R^2	0.126	0.124	0.136	0.319	0.203
Clusters	1,085	1,085	1,085	1,085	1,085
N	61,410	61,410	61,410	61,410	61,410

Notes: * p<0.05, ** p<0.01, *** p<0.001. The dependent variable contains the distance traveled including zeroes aggregated to the person-day level. The coefficients show the ATE in kilometers. “Heat day” and “Cold day” capture the positive deviations of daily temperature above 25 (heat) and below 7 (cold) degrees Celsius. Standard errors (in parentheses) are clustered at the participant level. All regressions include person and date fixed effects, as well as a dummy variable indicating days following the receipt of a negative travel budget in the mobility report.

Including weather controls neither improves the precision nor alters the magnitude of the treatment effect estimates. While cold temperatures show minimal impact on daily transport behavior, precipitation significantly decreases active mode distances, and heat reduces daily distances traveled by 4.5 km per degree above 25 degrees. Heat also reduces the use of public transport and walking. Alternative specifications, such as dummy variables for extreme temperatures, similarly did not affect the significance or magnitude of the treatment effects. Controlling for weather in regressions on external costs, as shown in Table A.6, likewise yields no improvements in precision.

Table A.6: ATE of external costs controlling for weather

	(1) Total Ext.	(2) Environm. Ext.	(3) Congest. Ext.	(4) Health Benefits	(5) Health Costs	(6) Accid. Ext.
Treated	-0.210*	-0.048	-0.068***	0.106***	-0.042	0.054
	(0.091)	(0.025)	(0.020)	(0.026)	(0.032)	(0.029)
Heat day	-0.401	-0.088	-0.019	-0.066	-0.190	-0.169*
	(0.237)	(0.046)	(0.028)	(0.046)	(0.104)	(0.085)
Cold day	-0.005	-0.001	-0.002	0.000	-0.002	0.001
	(0.009)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Precipitation	0.001	-0.001	-0.000	-0.009***	-0.000	-0.007***
	(0.009)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
adj. R^2	0.116	0.127	0.114	0.324	0.115	0.222
Clusters	1,085	1,085	1,085	1,085	1,085	1,085
N	61,410	61,410	61,410	61,410	61,410	61,410

Notes: * p<0.05, ** p<0.01, *** p<0.001. The dependent variable is the external cost of transport (in CHF) aggregated to the person-day level. "Heat day" and "Cold day" capture the positive deviations of daily temperature above 25 (heat) and below 7 (cold) degrees Celsius. Standard errors (in parentheses) are clustered at the participant level. All regressions include person and date fixed effects, as well as a dummy variable indicating days following the receipt of a negative travel budget in the mobility report.

A.5 Multivariate regressions

Table A.7: Multivariate interactions: External costs

	(1) Environm. Ext.	(2) Congest. Ext.	(3) Health Benefits	(4) Health Costs	(5) Accid. Ext.
Treated	0.047	0.078	0.069	0.054	0.067
	(0.068)	(0.089)	(0.086)	(0.079)	(0.060)
Treated x Male=1	-0.046	-0.078	-0.026	-0.057	-0.030
	(0.036)	(0.055)	(0.039)	(0.042)	(0.030)
Treated x Age>=50	-0.065	-0.076	-0.042	-0.074	-0.069*
	(0.040)	(0.055)	(0.042)	(0.045)	(0.030)
Treated x Tertiary educ.=1	0.010	0.006	-0.036	0.016	-0.028
	(0.040)	(0.053)	(0.043)	(0.046)	(0.034)
Treated x HH size<3	-0.012	-0.031	0.014	-0.016	0.002
	(0.040)	(0.055)	(0.044)	(0.046)	(0.034)
Treated x French=1	0.036	-0.014	-0.036	0.068	0.039
	(0.056)	(0.077)	(0.053)	(0.059)	(0.044)
Treated x Urban=1	-0.028	-0.082	0.090*	-0.054	0.010
	(0.036)	(0.048)	(0.039)	(0.041)	(0.030)
Treated x PT reduction=1	0.007	0.015	0.017	0.035	0.020
	(0.045)	(0.058)	(0.051)	(0.051)	(0.036)
Treated x S-pedelec=1	-0.045	-0.084	0.065	-0.043	0.054
	(0.040)	(0.055)	(0.038)	(0.047)	(0.032)
Pseudo R^2	0.104	0.131	0.219	0.114	0.130
Clusters	1,085	1,085	1,085	1,085	1,085
N	61,410	61,410	61,410	61,410	61,410

Notes: * p<0.05, ** p<0.01, *** p<0.001. The dependent variable is the external cost of transport (in CHF) aggregated to the person-day level. All dimensions include one omitted category. *Treated* is thus associated with an observation that has a zero for all included dummies. The coefficients were estimated using a PPML model, and the results show proportional effects. Standard errors (in parentheses) are clustered at the participant level. All regressions include date and person fixed effects, as well as a dummy variable indicating days following the receipt of a negative travel budget in the mobility report.