Pigovian Transport Pricing in Practice

May 16, 2023

Abstract

Pigovian transport pricing is implemented in a large-scale field experiment in urban areas of Switzerland. The pricing varies across time, space and mode of transport. One third of the participants is given a financial incentive to reduce their external costs of transport, whereas others are provided information only or serve as a control group. The pricing treatment causes a significant reduction in the external costs of transport. This reduction is a consequence of mode substitution and a shift in departure times. The effect of providing information in the absence of pricing is also statistically significant, implying that information and pricing each play an important role in explaining the total effect.

Keywords: Transport pricing; Pigovian taxation; mobility; external costs; congestion; tracking.

JEL Codes: H23, H31, I18, Q52, Q54, R41, R48

1 Introduction

Transport systems face multiple challenges. In many cities around the world, drivers lose over 100 hours per year due to traffic congestion (INRIX, 2020). Public transport can help reduce congestion (Anderson, 2014), but also faces crowding problems. Increasing the capacity of private and public transport faces physical limitations and high costs due to competition with other land use. Furthermore, greater road capacity induces demand and does not alleviate congestion (Duranton and Turner, 2011). The transport sector is also among the largest contributors of local air pollution (EEA, 2019) and greenhouse gas emissions (Creutzig et al., 2015), which have remained roughly constant during the past 30 years, as gains in efficiency have been neutralized by increases in distance traveled (IEA, 2020).

Congestion, climate damages and health effects constitute the most important external costs of transport. Whereas the private costs of transport, such as the purchase of fuel or a transport pass, have been shown to influence individual transport choices (Oum, Waters and Yong, 1992; Goodwin, Dargay and Hanly, 2004; Vrtic et al., 2008), the external costs are borne by society at large and are typically not reflected in the decision about where, when and how to travel.[1] This large-scale market failure is the normative motivation for policy interventions in the transport domain. In this paper, we implement a multi-modal Pigovian transport pricing scheme based on the full marginal social costs of transport and estimate its effects on individual transport choices.

Our study employs a randomized controlled trial (RCT) design, which allows for unbiased estimates of treatment effects. The sample consists of people living in urban areas of the German- and French-speaking regions of Switzerland. The pricing affects all modes and is implemented by providing the participants with a personalized budget[2], from which the external costs of their transport choices are subtracted. This treatment causes a reduction in the external costs of transport of around 5.2%. If we relate this to the (total) price increase implied by the Pigovian pricing scheme of 16.4%, we arrive at an elasticity of -0.32. The reduction in the external costs is due to a shift in transport mode and departure time, but not a reduction in the overall distance traveled. Car owners, people living in rural areas, and the young (under 30 years old) respond by more than the average to the pricing treatment, whereas French speakers respond less. We further find that

[1] The third category of transportation costs are the infrastructure costs, which are fixed and can be paid by user fees or general subsidies. For a recent discussion of the definition of the external costs of transport, see CE Delft (2019).

[2] The budget for the intervention phase is based on the observed external costs from the initial four-week observation phase.

the effect is twice as large for those that understood the experiment and/or had prior knowledge about the external costs of transport.

To differentiate the pricing effect from a pure information effect, the experiment includes a second treatment arm in which the participants are provided with the same information about the external costs of transport as the pricing group, but without having to pay anything. The results suggest an effect of information *per se*, which varies strongly over participants' intrinsic values. Last, the differential effect between the pricing and information groups can be interpreted as the causal effect of adding a price to existing information. This "pricing only" effect is also statistically significant. Our results thus imply that information and monetary incentives each play an important and separate role in explaining the total effect of the pricing treatment.

The external costs of transport have been, for the most part, addressed by "command-and-control" policies such as speed limits (Van Benthem, 2015), fuel standards (Portney et al., 2003), license-plate restrictions (Davis, 2017) or high occupancy lanes (Bento et al., 2014). Theoretically speaking, however, price instruments reflecting the external costs of transport are a more efficient means of regulation as they allow people to retain high-utility trips while reducing those that they view as less important.

The most prevalent examples of price-based instruments in the transport sector are fuel taxes, road tolls and registration fees. However, they are usually imposed to recover the cost of road construction and thus typically do not reflect the full external costs of transport (Parry and Small, 2005; Parry, Walls and Harrington, 2007). Road congestion charges can act as an effective way to internalize some of the external costs of driving (Small, 2008), and several cities have introduced fees for driving into the city center at certain times. However, these fees tend to be fixed and can therefore not fully address the time-varying nature of congestion. Furthermore, these policies generally target only one transport mode, which decreases efficiency and raises concerns about equity within the transport sector.

The theoretical foundations for road pricing were laid by Pigou (1920) and Knight (1924). Vickrey (1963) showed that with optimal congestion pricing, tolls must match the severity of congestion, and vary by time of day, location, type of vehicle and current conditions. The methods for road pricing can be categorized by their level of internalizing the external costs. In first-best pricing, the marginal external cost is charged to the user. In this case, both the charging mechanism and the amount charged need to be optimal (Verhoef, 2000). In second-best, the pricing mechanism is guided by the principle of marginal external costs, but the implemented scheme is simplified (Small, Verhoef and Lindsey, 2007). For a review of studies that applied economic incentives in the transport context, see Dixit et al. (2017).

Eliasson (2021) provides an overview of the state of the art in transport pricing with regard to necessity, efficacy, equity, and feasibility while Beheshtian et al. (2020) propose a multi-modal network management scheme for congested transportation systems based on insights from efficient electricity market mechanisms. Previous empirical research includes computations of the aggregate effects of the congestion charges that were introduced in Singapore (Agarwal and Koo, 2016), London (Leape, 2006), Stockholm (Eliasson et al., 2009) and Gothenburg (Börjesson and Kristoffersson, 2018). Field experiments in Denmark and Australia installed GPS receivers in vehicles and drivers were then exposed to different peak and off-peak pricing schemes (Nielsen, 2004; Martin and Thornton, 2017). Commuters in Singapore responded to rewards and social comparisons by shifting departure times when using public transit (Pluntke and Prabhakar, 2013), and quasi-experimental evidence of the congestion charges in Norway and Milan suggests that they were effective in reducing congestion and air pollution (Isaksen and Johansen, 2021; Gibson and Carnovale, 2015). Yang, Purevjav and Li (2020) estimate the impact of an optimal road congestion charge for Beijing and find an increase in both peak time travel speeds and social welfare. There are two other RCTs involving financial incentives that we are aware of. One is by Rosenfield, Attanucci and Zhao (2020), who carry out an experiment involving 2,000 employees at the Massachussetts Institute of Technology. They use three treatment arms, but none of them led to statistically significant effects. The other is by Kreindler (2018), who examines the effect of a departure time charge and a zonal price on drivers in Bangalore and derives highly significant treatment effects using a smartphone app similar to ours.

Most existing studies focused on a single mode of transport and could therefore not identify modal shifts. To detect the full impact of transport pricing, observation (and ideally pricing) of all modes is necessary, including of non-motorized modes (Tirachini and Hensher, 2012) An example of a previous multi-modal study is the "Spitsmijden" experiment in the Netherlands, in which commuters responded to financial and in-kind rewards by shifting departure times, switching to other modes of transport and by working from home (Ben-Elia and Ettema, 2011). Their monetary rewards were comparable to the average external cost in our experiment. Our study is similar in spirit but based on a larger sample and, importantly, uses a control group to absorb time-varying factors that may be correlated with the treatment. This makes MOBIS, to the best of our knowledge, the first multi-modal RCT of a pricing intervention in the transport context.

Because of concerns related to social acceptability, behavioral change could also be achieved by means of non-financial interventions, which may be easier to implement than prices or taxes. A number of studies have investigated the effect of non-financial interventions in the transport sector (see Möser and Bamberg, 2008, for a review), and some recent papers have used tracking apps to test the effect of informational interventions, however, these are based on small samples (Maerivoet et al., 2012; Carreras et al., 2012; Bothos et al., 2014; Jariyasunant et al., 2015). Among the few information-based RCTs carried out in the transport context are Cellina et al. (2019) and (Kristal and Whillans, 2020), both of which did not find any effects of non-financial incentives.

Our paper makes several contributions to the literature of transport economics. First, we provide a methodology to compute the marginal external costs of transport as they vary across time and space and apply it to a sample observed in real time. This allows for the computation of a multi-modal Pigovian pricing scheme, which provides a useful benchmark for simplified versions of transport pricing. Second, by implementing this pricing scheme in a representative sample of the population living in large urban agglomerations, we obtain credible information about the short-run behavioral response to multi-modal transport pricing, including modal substitution. Third, because we apply an information-only treatment within the same experiment, our experiment contributes to our understanding of the relative importance of information-based and monetary incentives in the transport domain.

The next sections provide more background about the experimental setup and the methodology used to compute the external costs of transport. Section 4 describes the methodology, section 5 contains the results, and section 6 concludes.

2 The MOBIS experiment

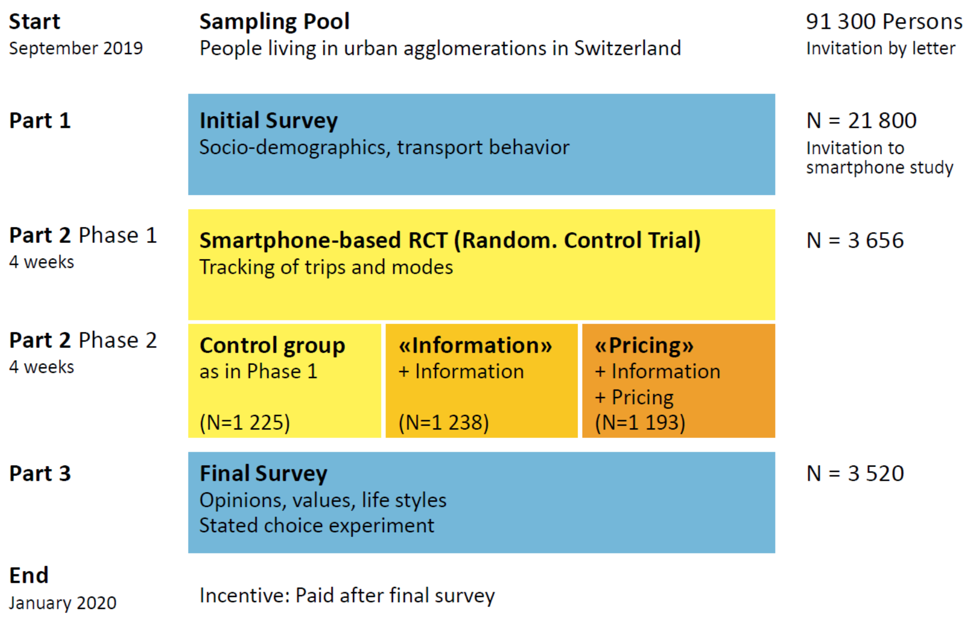
2.1 Study design and sampling

The sample for the Mobility in Switzerland (MOBIS) project was recruited among individuals living in urban areas in the German- and French-speaking regions of Switzerland. The study participants agreed to having their daily travel tracked with a smartphone app (see below) over a period of 8 weeks. In return for taking part in the study, all participants were offered CHF 100, which they received at the end of the project.[3] The participants in the pricing treatment were paid out their remaining budget (if nonzero) in addition to the CHF 100 participation incentive.

Figure 1 provides an overview of the study design. We contacted a representative sample of 91,000 people by mail and invited them to participate in the study. The

[3]In September 2019, one Swiss Franc (CHF) corresponded to 0.92 Euro and 1.01 US Dollars.

Figure 1: Design of the MOBIS experiment



addresses were randomly selected and provided by the Federal Office of Statistics, which maintains a comprehensive registry of inhabitants as well as from a private vendor.[4] The first step consisted of an initial online survey, which was completed by 21,800 respondents. It contained questions about travel behavior and socio-demographics and served as a screening mechanism for the RCT. To be invited for the tracking part of the study, respondents had to use a car on at least 2 days a week and they could not be professional drivers. Around 11,000 respondents from the initial survey qualified for the tracking study, and 5,466 registered. However, not everyone who registered actually started tracking, and some participants dropped out at the beginning of the study. A total of 3,656 participants completed the RCT.

Because of the "double" self-selection (first into the survey and then into the tracking part of the study) and the driving requirement, a careful look at the composition of our sample is warranted. Table 1 shows the socio-demographic characteristics of participants in the introduction survey and the RCT compared to the Mobility and Transport Micro-census (MTMC), which is a representative travel diary survey of the Swiss population undertaken by the Federal Office of Statistics and the Federal Office of Spatial Development (2017). To provide a meaningful comparison, we restrict the MTMC sample to the same age range (18-65) and geographic area as our study (postcodes identified as being

[4]We were provided with 60,000 addresses from the Federal Office of Statistics at no charge. When it became clear that this would not be sufficient to recruit the required number of participants, we purchased another 31,000 addresses from a private marketing firm.

part of major urban agglomerations). The respondents of the MOBIS introduction survey are overall quite similar to the MTMC population. The largest differences are in terms of the share of young adults aged 18-25 (20.1% vs. 14.3%), education (47.5% vs. 38.7% with tertiary education) and being Swiss nationals (including additional nationalities).[5] Employment, gender, household size, income, language and degree of urbanisation[6] are similar.

The tracking sample has a slightly higher employment rate, more students, and fewer one-person households than both the Intro survey and the MTMC samples, but is similar along most other socio-demographic characteristics. The degree of urbanisation has a higher share of "intermediate", which is most likely due to the car driving requirement for participation in the study.Due to our procedure of unconditional random assignment, the socio-demographic variables are evenly distributed (but not identical) across the three RCT groups.

Mobility tool ownership may be an important determinant for responding to transport pricing. The percentage of people that do not have access to a car is lower in the RCT sample than in the MTCM (even though both are in the single digits), most likely because we conditioned participation on regularly driving. For the same reason, the share of people who commute by public transport (and thus have a full PT subscription) is somewhat lower in our sample than in the MTMC (25% vs. 35%). However, the share of people who hardly ever use public transport and therefore do not have a half-fare subscription is similar across the samples. We discuss the implications of our sample selection procedure for external validity in section 6.

More details about the design and implementation of the experiment can be found in Appendix B.

2.2 Tracking app

The participants in the tracking study agreed to download the tracking app "Catch-My-Day" on their smartphones. Catch-My-Day is a location tracker for iOS and Android, which uses the location services of the respective operating system. The GPS tracks are

[5] We believe this due to the fact that our recruitment was based on letters and online surveys, whereas the MTMC is based on targeted telephone interviews. If necessary, the MTMC uses translators. In contrast, people who are not fluent in English or one of the national languages likely disregarded our invitation.

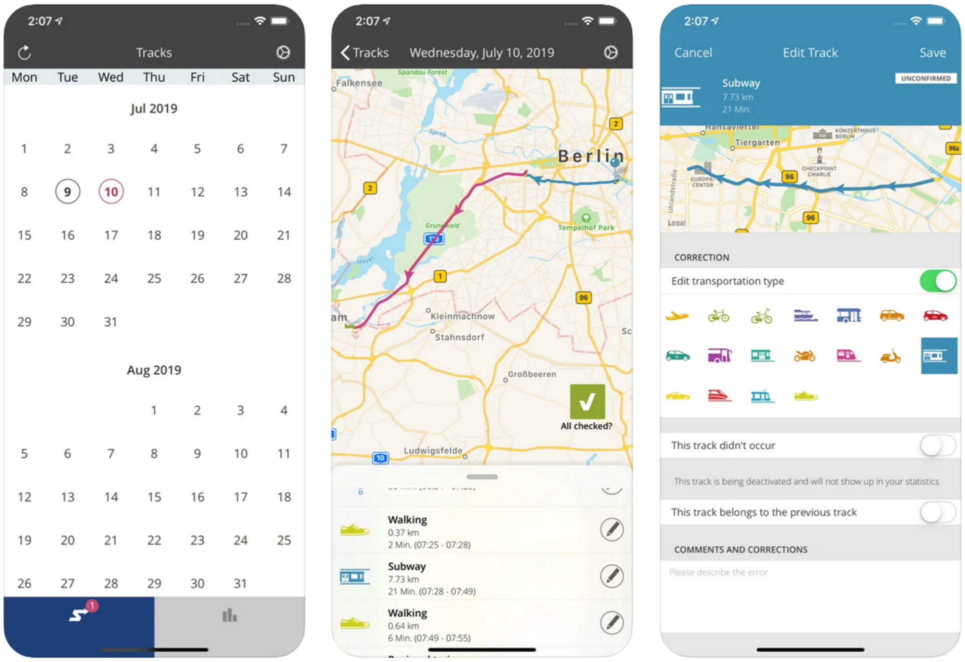
[6] This variable is constructed by allocating participants' home postcodes one of three degrees of urbanization: urban, intermediate, rural. These definitions are based on the Swiss Federal Statistical Office's definitions, which is partly based on the accessibility of road and public transport infrastructure (Federal Statistical Office, 2017)

Table 1: Demographic information for the MOBIS sample

Variable	Level	Intro	Tracking			MTMC
			Control	Info	Pricing	
Age	[18, 25]	20.1	18.4	19.9	19.8	14.3
	(25, 35]	19.4	18.0	18.7	16.9	21.4
	(35, 45]	19.9	22.3	21.1	24.6	22.6
	(45, 55]	21.6	22.9	23.9	22.5	23.7
	(55, 65]	19.0	18.4	16.4	16.2	17.9
Education	Mandatory	9.2	8.0	5.2	6.9	13.8
	Secondary	43.3	47.4	49.3	48.5	47.5
	Higher	47.5	44.5	45.5	44.6	38.7
Employment	Employed	68.7	73.5	72.0	70.7	68.8
	Self-employed	7.3	6.0	5.2	7.3	8.8
	Apprentice	1.9	1.9	1.6	1.6	2.2
	Unemployed	4.4	3.3	3.8	4.7	3.9
	Student	9.3	7.8	8.8	7.9	3.0
	Retired	2.5	2.6	2.0	2.3	3.6
	Other	5.9	4.8	6.5	5.5	9.7
Gender	Male	48.9	50.0	49.9	49.1	49.4
	Female	51.1	50.0	50.1	50.9	50.6
Household size	1	15.5	11.2	11.3	12.1	18.3
	2	31.7	30.0	31.0	28.6	32.0
	3	20.5	23.0	21.6	19.9	19.9
	4	23.6	25.5	28.0	29.9	20.7
	5 or more	8.6	10.3	8.2	9.5	9.1
Income	4,000 CHF or less	12.2	6.6	8.3	7.2	8.8
	4,001 - 8,000 CHF	29.4	31.3	29.9	27.5	31.4
	8,001 - 12,000 CHF	24.5	27.8	30.0	30.4	24.6
	12,001 - 16,000 CHF	12.1	15.6	13.6	14.3	11.7
	More than 16,000 CHF	8.0	9.6	9.6	10.4	8.4
	Prefer not to say	13.8	9.1	8.6	10.1	5.8
	Don't know					9.2
Language	German	62.7	66.5	65.5	66.6	69.5
	French	28.6	25.7	26.2	26.5	26.5
	Italian					4.0
	English	8.7	7.8	8.3	7.0	
Nationality	Switzerland	78.1	81.3	80.6	82.5	69.5
	Other	21.9	18.7	19.4	17.5	30.5
Area	Urban	75.0	63.7	64.6	64.4	77.4
	Intermediate	18.1	27.9	26.9	28.0	16.6
	Rural	6.8	8.3	8.5	7.7	6.0
Access to car	Yes	61.0	87.3	88.0	87.8	69.7
	Sometimes	15.5	11.7	10.6	11.3	22.7
	No	23.5	1.0	1.3	0.9	7.5
Full PT subscription	Yes	37.2	21.9	25.2	25.7	34.5
Half fare PT subscription	Yes	47.6	49.2	49.1	47.7	37.6
No PT subscription	Yes	26.0	33.7	32.6	34.0	37.9
Access to bicycle	Yes	68.5	72.8	72.1	69.8	70.1
	Sometimes	4.1	4.5	5.6	4.0	8.8
	No	27.4	22.7	22.2	26.3	21.1
N		20,783	1,174	1,187	1,134	21,399

Notes: Descriptive statistics shown for the MOBIS introduction survey sample, the MOBIS tracking sample (which is a subset of the former), and the weighted Swiss Mobility and Transport Microcensus 2015 (MTMC) sample. All samples restricted to 18 to 65 year olds, with the MTMC sample additionally restricted to respondents living in municipalities present in the MOBIS introduction survey sample.

Figure 2: The Catch-my-Day interface



Notes: From left to right: 1) Calendar home page. 2) Daily view showing recorded trips. 3) Editing the mode of a selected trip.

stored on the phone and uploaded to the Motiontag analytics platform, where trip stages are identified and travel modes and activities are imputed based on a machine learning algorithm. For each stage, the associated external costs of transport were computed based on cost factors published by the Swiss Government (see section 3). Participants were able to review and correct the mode assignment manually.

Figure 2 shows three interfaces of the app. Catch-my-Day provides a best guess of the travel mode for each stage. The participants could confirm this detected mode or correct it. This confirm-correct procedure was optional and participants were informed that this was possible and would be appreciated.[7] Around 29% of the stages were confirmed by the participants. The database stores both their correction and the original algorithmic imputation. The possibility for mode correction increases the accuracy of the mode detection, but it also introduces a scope for "gaming" the experiment. We return to this issue in section 6.

The following modes are detected the by Catch-my-Day app: Airplane, bicycle, bus,

[7] In recent years, state-of-the-art machine learning algorithms for mode and activity detection have achieved accuracy rates of over 90%, depending on the approach (Wu, Yang and Jing, 2016; Nikolic and Bierlaire, 2017). Hence, we made validation of the trip purpose and mode optional for participants, in order to not increase the response burden excessively over the 8 weeks.

car, ferry, train (local, regional and long-distance), tram and walk. In addition, users could select the following modes as a correction: Boat, car sharing, gondola, motorbike/scooter, taxi/Uber. E-bikes and E-Scooters were neither detectable nor selectable.

The mode detection provided by the tracking app was a key component of the MOBIS study. To the best of our knowledge, this is the first study to incentivise changes in mobility behavior based on the output of a mode detection algorithm. The algorithm worked very well and achieved an overall accuracy of over 90% (see Tables B.1 and B.2) in the Appendix).

2.3 Treatments

The recruitment took place on a rolling basis between August and November 2019. Once a participant registered their first track on the app, they automatically became part of the RCT sample. The participants did not know at this stage that they were part of an experiment. The RCT consisted of 4 weeks of observation for all participants, followed by another 4 weeks of one of two treatments.[8] Assignment to the treatment and control groups was fully randomized without any form of stratification. During the observation period, participants were presented with a weekly summary of their travel behavior by mode of transport, including duration, distance and number of trips.

On study day 29, the participants assigned to the "Information only" and "Pricing" treatments received an e-mail that informed them about the external costs of transport, how they are computed and monetized and what they could do to reduce them. The e-mail contained a link to a table with per-km monetized costs by mode. The congestion cost was framed as an example, as it varied by time and place. To complement this ex-ante price information and to provide the participants with an idea about their individual level of external costs (i.e., including congestion), they were also shown a personalized summary of the previous week.[9] For the remainder of the treatment period, the participants were presented with weekly summaries such that they could observe changes in their external

[8]The study concluded just before the onset of the COVID-19 pandemic at the beginning of 2020. Some of the participants agreed to re-start tracking, as part of an effort to study travel patterns in response to COVID-19 policies. Preliminary results of this ongoing study have been reported in Molloy et al. (2020).

[9]To provide participants with ex-ante information about the congestion costs for particular trips was infeasible within the project budget as this would have required a lot of additional programming and presumably also a different app. However, it is not clear to what extent the participants would have actually relied on such detailed information. Furthermore, the internal part of congestion costs are experienced personally, such that participants have an idea about the expected congestion in their area. We believe that combining ex-ante averages with ex-post individualized numbers is a reasonable compromise that sends a price signal without overly taxing participants' attention.

costs. The external costs were always presented by mode of transport and by type of cost (health, climate and congestion).

The participants assigned to the "Pricing" treatment received the same information about the external costs as the "Information" group, but in addition were given a budget from which the external costs of transport were deducted. The participants were informed on day 29 that as of now, their budget would be used to pay for the external costs caused by their travel, and that any money left over in their account at the end of the study was theirs to keep. This individualized budget was computed based on each participants' external costs during the observation period, plus a 20% buffer to allow for the possibility that some participants had to increase their external costs of transport for idiosyncratic reasons.[10] This treatment thus simulated transport pricing based on the monetized marginal external costs of transport.

The weekly reports were comprised of modular panels, as shown in Figure 3. The introduction and distance by mode panels were presented to all participants in both study phases. The external cost and chart explanation panels were shown to the information and pricing groups in the treatment phase, and the remaining budget panel (middle module on the left) only to the pricing group during the treatment phase. Due to the rolling start of the experiment, participants received these reports on different days of the week. Participants in the control group continued to received a weekly email with their kilometers traveled per mode throughout the experiment. Once the RCT was concluded, all participants were informed about having taken part in a research experiment.[11]

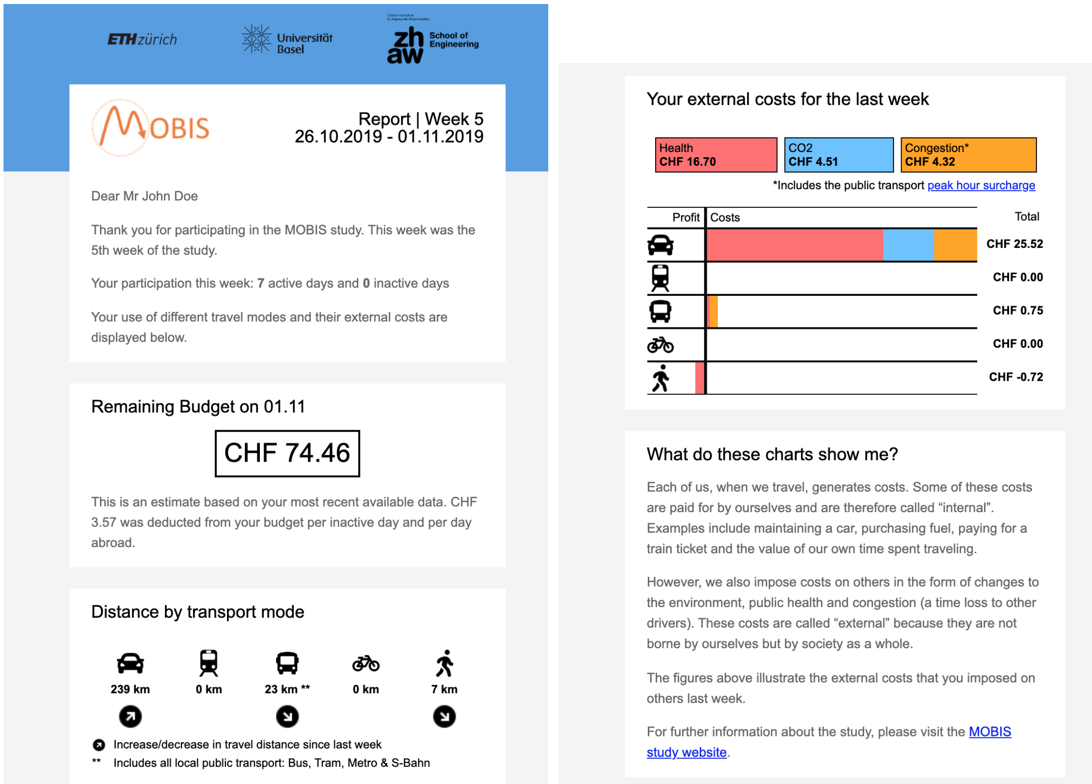
In principle, we could have used any pricing scheme and estimated the participants response to it. We chose the Pigovian rate (i.e., the marginal external costs) for three reasons. First, internalizing the external costs of transport can be motivated and explained to people on normative grounds. The "information only" treatment could thus be interpreted as providing information on societal costs about which the participants were likely not perfectly informed. In contrast, introducing a price unrelated to the external costs would not contain useful information (other than the price signal itself) and thus not lead to behavioral change via an altruistic motive. Second, using the Pigovian rate and estimating people's response to it serves as a policy benchmark.[12] If larger (smaller)

[10]The average budget was CHF 144, but for some participants it exceeded CHF 700. Note that once participants reached the treatment phase, they could no longer correct the modes during the observation phase in order to prevent a strategic increase of the mobility budget.

[11]This procedure was pre-approved by ETH's Institutional Review Board.

[12]Technically speaking, the Pigovian rate is the marginal social damage *at the social optimum*, such that the pricing implemented in the experiment likely deviates from the true Pigovian tax. If such a scheme were implemented in practice, however, one would need to monitor the external costs anyway and update the scheme from time to time, such that the social optimum would be reached iteratively.

Figure 3: Weekly reports by e-mail



Notes: The participants in the control group received only the report on the left, but without the middle module titled “Remaining Budget”. The participants in the information group additionally received the message on the right. The participants in the pricing group received all modules.

responses are required, the policy maker can choose to exceed (stay below) this rate, but we believe that knowing the level and people’s response to the first-best transport price is useful information. Last but not least, the use of the Pigovian rate was a condition imposed by one of the federal agencies that co-funded the project.

3 The external costs of transport

The health, emissions, noise and congestion costs of the mobility behavior were computed on the recorded daily trips using an automated data pipeline. Additionally, data collected from the online introduction survey was incorporated into the data processing pipeline to improve the imputation.

Table 2: Monetization of externalities

Externality	Type	Value	Unit
CO_2	Climate Costs	136.08	CHF/ton
PM_{10}	Rural	515,497	CHF/ton
	Urban	1,358,461	CHF/ton
NO_x	Regional	7,109	CHF/ton
VTTS	National	25.77	CHF/h

Source: Federal Roads Office - ASTRA (2017), updated for 2019; the value of travel time savings (VTTS) is the scaled nominal wage rate.

3.1 Costs associated with driving

For the calculation of external costs in private road transport, the recorded GPS tracks were aligned to the road network using Graphhopper (Karich and Schröder, 2014) and processed using modules developed on top of the MATSim framework to calculate the external costs of congestion and emissions. The emissions factors were taken from the HBEFA database (version 3.3), and applied using the MATSim emissions module (Hülsmann et al., 2011; Kickhöfer et al., 2013). For congestion, an average marginal cost approach incorporating spillback effects and flow congestion was applied, based on the work of Kaddoura (2015). These modules returned quantities of the externalities in grams (for emissions) and seconds of caused delay (for congestion) for road transport, which were then converted to monetary costs using the values in Table 2. The computation of the externalities associated with private motorized transport is shown schematically in Figure A.1 in the appendix.

3.2 Costs of public and active transport

For modes other than driving, the per-km values presented in Table 3 were applied to the recorded length of the trip. The health effects include accident costs (most of which are external to the people involved due to coverage by the Swiss health care system), but also the external portion of health benefits in the form of a reduced mortality and morbidity as a consequence of physical activity (Götschi, Garrard and Giles-Corti, 2016). Whereas walking is associated with net external benefits, the external accident costs outweigh the external health benefits from cycling, such that bicycling is associated with small net

Table 3: Per-km monetary costs (in CHF) used in the MOBIS experiment

				Total	
Mode	Congestion	CO_2	Health	w/o Congestion	incl. Congestion
Car	0.0332	0.0258	0.0781	0.1039	0.1371
Train	0 / 0.1	0.00007	0.0141	0.0141	0.1141
Tram	0 / 0.1		0.0141	0.0141	0.1141
Bus	0 / 0.1	0.0144	0.0710	0.0854	0.1854
Bicycle			0.07	0.07	
Walk			-0.11	-0.11	

Notes: The values for public and active transport are based on NISTRA (Federal Roads Office - ASTRA, 2017). Congestion costs for public transport were only applied for congested links (see text). Negative costs indicate an external benefit. The external costs of driving vary over time and space and were computed within MATSim (see Fig. A.1). The values shown are the average per-km costs from the tracking data.

external costs in the experiment.[13]

The marginal social cost of public transport (in terms of pollution and noise) decreases as the occupancy rate increases. On the other hand, crowding affects willingness to pay and can be seen as a form of congestion in public transport, and delay in some circumstances (Tirachini, Hensher and Rose, 2013). Crowding effects are extremely heterogeneous, both spatially and temporally. Even in peak hour, crowding can be restricted to particular transit lines during very short periods (Zurich Public Transport, 2017). We felt that it would be unreasonable to distribute the crowding effects in an aggregate measure across peak hour travelers in a specific public transit region. Additionally, for each public transport operator, data would have to be collected separately and collated as it is not available on a national level. As a practical solution, a zonal peak-hour surcharge pricing scheme was developed for the national public transport network as a form of second-best pricing. Throughout the experiment, participants had access to an interactive map which showed them where and when the pricing scheme applied. The peak-hour pricing surcharge of 0.10 CHF/km was applied to transit stages between any two zones which experienced peak hour demand. The peak hour windows and the affected zone-pairs were determined using the MATSim scenario output for Switzerland (Bösch, Müller and Ciari, 2016). The peak windows were set as 7am to 9am and 5pm to 7pm and not adapted for regional variation in working patterns. Municipality pairs were priced if the

[13] Most of the positive health effects are private in the form of lower morbidity and mortality and at least partly internalized by cyclists (Götschi and Hintermann, 2014).

maximum hourly transit trip count during peak hour was greater than three times the average hourly transit trip count during the daily off-peak period (9am - 5pm) for that pair. A municipality could also be paired with itself if the above criteria were met and the direction of the peak hour flow was not considered. If the trip was partially in both the peak and off-peak periods, only the proportion of the travel duration that overlapped with the peak period was charged.

4 Regression framework

Given that we randomize the treatment and thus no self-selection or endogeneity issues need to be addressed, the average treatment effect (ATE) can be estimated by comparing means between treated and control observations. We aggregate the data to the person-day level and estimate the ATE using the following regression:

$$Y_{its} = c_0 + \alpha^P \cdot DiD^P_{its} + \alpha^I \cdot DiD^I_{its} + \mu_i + \mu_t + \mu_s + \epsilon_{its} \tag{1}$$

The dependent variable is the outcome of interest for person $i \in (1, ..., N)$ on calendar day $t \in (1, ..., T)$ and day of study $s \in (1, ..., 56)$. The main outcome of interest is the total external cost (in CHF per day), but we also run regressions where the dependent variable is the external cost along a particular dimension (health, climate and congestion), the distance traveled in total or by mode or the time of departure.

The two difference-in-differences terms DiD^P_{its} and DiD^I_{its} are the products of treatment group and treatment period dummies and are equal to one if the pricing (P) and information treatment (I), respectively, are active for person i on a given day, and zero otherwise. The treatment starts on the 29th day of the experiment. Due to the rolling recruitment, the beginning of the experiment varies by person.

To control for unobserved heterogeneity, we include fixed effects on the person (μ_i), calendar day (μ_t) and day-of-study (μ_s) level. The calendar day FE capture common shocks that affect travel (and thus the associated external costs) for everyone in Switzerland, e.g., due to a national holiday or a sports event. The day-of-study FE account for the possibility that respondents may respond differently to the treatment over time. The combination of day-of-study and calendar day FE implies that the treatment effect is computed by comparing participants in the treatment and control groups that started the experiment on the same day. Finally, the error term ϵ_{its} has an expected mean of zero and a variance of σ^2. We allow for a correlation of the error within participants.

The ATE of pricing plus information is given by the coefficient estimate α^P; the ATE for information only is given by α^I; and the ATE of adding pricing to information is their difference, $\alpha^P - \alpha^I$. We emphasize that $\alpha^P - \alpha^I$ could also be computed by running a DiD analysis on the pricing group while using the information group as the control. It is therefore a causal ATE in its own right, rather than simply a difference between two coefficients.

Due to the random assignment, we do not need to control for any covariates as they are expected to affect the treatment and control groups equally. However, because our sample is finite and weather information is an important predictor especially for active transport, we enrich our tracking data with temperature and precipitation data from MeteoSwiss provided on a 1 x 1 km grid.[14] This could reduce the noise in the regression and thus increase the precision of our estimates, but at the cost of introducing a parametric assumption. The weather variables are assigned separately for each recorded trip based on the nearest weather station. To allow for a nonlinear effect of temperature on travel choices, we define the level of "Heat" and and "Cold" for an observed trip j on day t relative to threshold values:

$$Heat_{jt} \equiv \max\{t_{jt}^{max} - 25,\, 0\} \tag{2}$$

$$Cold_{jt} \equiv \max\{10 - t_{jt}^{min},\, 0\} \tag{3}$$

The variables t_{jt}^{max} and t_{jt}^{min} refer to the daily maximum and minimum temperature, respectively, recorded in degrees Celsius at the weather station closest to the departure location for trip j. We compute the average of the heat, cold and precipitation values across all trips taken by person i on day t and add them as linear control variables to (1).

To investigate potential differences of the treatment effect along major socio-economic variables, we further interact the DiD terms with categorical variables denoting, e.g., gender or income groups.

For the regressions that use external costs as the dependent variable, we estimate (1) in levels. Estimation in levels (rather than in logs) is necessary as the external benefit associated with walking renders a number of person-day observations negative. We then compute the proportional response by dividing the coefficients (which are in CHF) by the external costs generated during the observation period. For regressions in which the dependent variable is non-negative (e.g., distance traveled), we estimate proportional effects directly by using a Poisson Pseudo-Maximum Likelihood (PPML) model. This approach addresses the possible presence of heteroskedasticity, which can lead to a bias

[14]The data is provided by `www.meteoswiss.admin.ch`.

Table 4: Tracking summary statistics

Dimension	Outcome	Unit	Pre-treatment Control	Pre-treatment Info	Pre-treatment Pricing	Post-treatment Control	Post-treatment Info	Post-treatment Pricing
External costs	Total	CHF/d	4.51	4.60	4.71	4.24	4.27	4.25
			(5.66)	(5.63)	(5.80)	(5.39)	(5.53)	(5.41)
	Congestion	CHF/d	1.04	1.07	1.14	0.85	0.87	0.90
			(1.58)	(1.58)	(1.69)	(1.44)	(1.53)	(1.52)
	Climate	CHF/d	0.88	0.88	0.90	0.85	0.84	0.83
			(1.29)	(1.29)	(1.30)	(1.23)	(1.28)	(1.22)
	Health	CHF/d	2.59	2.64	2.67	2.53	2.56	2.52
			(3.53)	(3.53)	(3.60)	(3.47)	(3.56)	(3.46)
Private costs	Total	CHF/d	26.06	26.69	26.87	25.71	25.96	25.56
			(33.59)	(33.92)	(34.49)	(33.39)	(34.02)	(33.21)
Tracking	Distance	km/d	46.96	48.02	49.47	45.55	47.25	47.41
			(55.23)	(54.59)	(57.31)	(54.17)	(55.88)	(55.01)
	Duration	min/d	92.77	93.63	94.42	88.57	90.76	91.38
			(84.71)	(80.16)	(82.87)	(78.85)	(82.11)	(82.87)
	Total days	Nr.	23.14	23.30	23.17	22.99	23.00	22.81
			(3.73)	(3.58)	(3.87)	(3.93)	(4.04)	(4.19)
	Trips	Nr./day	4.71	4.74	4.76	4.53	4.49	4.55
			(3.02)	(3.00)	(2.96)	(2.81)	(2.79)	(2.83)

Notes: Average values per participant over the course of the study (SD in parentheses).

in log-linearized regressions, and the presence of zeroes in the data.[15]

5 Results

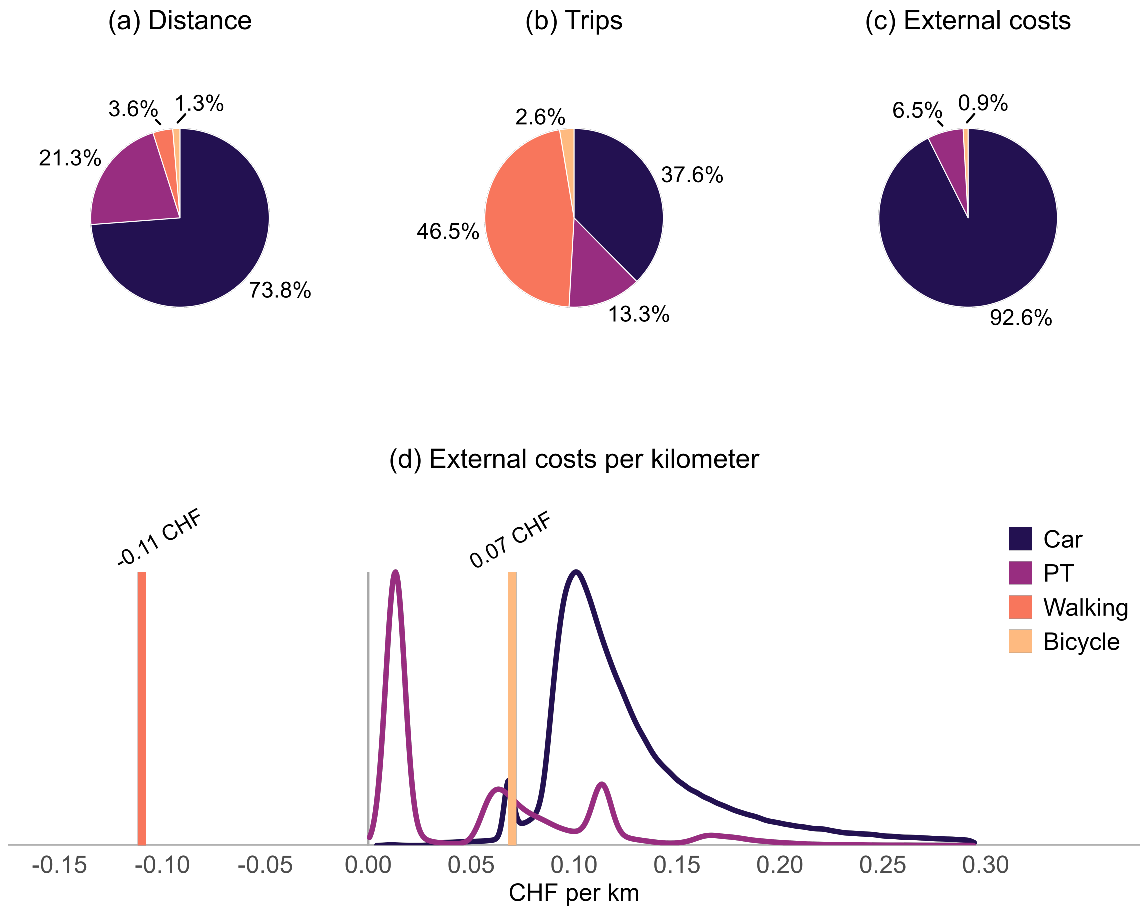
5.1 Tracking summary

Table 4 shows the summary statistics of the tracking data for overall travel, including distances, duration, external and private costs. Table A.1 in the Appendix shows the data separately for each mode.

Figure 4 displays the mode distribution recorded in MOBIS in terms of distance (panel a), number of trips (panel b) and external costs (panel c). Over 70% of the recorded distance is traveled by car, whereas the majority of stage counts correspond to walking. The vast majority of external costs is associated with driving. Panel d shows the constant per-km external costs for walking and bicycling, as well as the distribution of external per-km costs for driving and public transport, which vary over space and time due to congestion and crowding.

[15]For a discussion of the advantages of using a Poisson model in the presence of zeroes and heteroskedasticity, see Santos Silva and Tenreyro (2006). We use Stata's ppmlhdfe command that was developed by Correia, Guimarães and Zylkin (2019) and Correia, Guimarães and Zylkin (2020).

Figure 4: Mode distribution of distances, trips and external costs



Notes: For the external costs of walking and cycling, fixed values per person-km were used (see Tables 2-3). The external costs of car and public transport vary over time and space.

We observe a seasonal variation in the travel distance by mode, which translates to a negative trend in the external costs of transport. Including a control group allows us to absorb such trends. In contrast, if the treatment effect were estimated based on a before-vs.-after approach, as in most of the previous literature, the results would be biased towards a larger effect as the decrease in external costs over time would also be attributed to the treatment.

Before using the data for analysis, it was cleaned using some routine procedures that check for plausibility and remove obviously problematic tracking data (such as discrete jumps or physically impossible trips). We find a high variation on the first day (presumably due to technical issues with the app) and on day 29 (some people may have been slow to check their e-mails, which contained the treatment). We remove these two days from the analysis. We only included participants that delivered tracks on at least 4 days during both the treatment and the observation period (not counting days 1 and 29). In addition, we removed the data if one of the following was true: Average daily speed for car and PT above 100 km/h, above 40 km/h for bicycling and above 20 km/h for walking; or more than 500 km/day for car and PT, and more than 20 km/day for walking. If one of these limits was exceeded, we removed this person-day observation.

5.2 Average treatment effects

Table 5 shows the average treatment effect (ATE) on the external costs of travel in CHF per day. The first two columns report the results for the total external costs of transport, with and without controlling for the weather, whereas the next three pairs of columns contain the ATE on the external health, climate and congestion costs. About half of the reduction in external costs is due to a decrease in health costs, followed in magnitude by congestion and then climate costs. Including the weather does not significantly change the ATE.

Table 5: Average treatment effects on external costs

	Total ext. costs		Health costs		Climate costs		Congestion costs	
Pricing	-0.220**	-0.222**	-0.113*	-0.115**	-0.037*	-0.037*	-0.070**	-0.070**
	(0.071)	(0.070)	(0.044)	(0.044)	(0.016)	(0.016)	(0.022)	(0.022)
Information	-0.089	-0.094	-0.046	-0.049	-0.019	-0.020	-0.024	-0.024
	(0.068)	(0.068)	(0.043)	(0.042)	(0.016)	(0.016)	(0.021)	(0.021)
Difference	-0.130’	-0.128’	-0.067	-0.065	-0.017	-0.017	-0.046*	-0.046*
	(0.070)	(0.070)	(0.044)	(0.044)	(0.016)	(0.016)	(0.022)	(0.022)
Precipitation		0.002		-0.000		-0.000		0.002’
		(0.004)		(0.003)		(0.001)		(0.001)
Heat		0.191**		0.159**		0.058**		-0.026**
		(0.019)		(0.012)		(0.004)		(0.004)
Cold		-0.512**		-0.363**		-0.131**		-0.018
		(0.074)		(0.049)		(0.017)		(0.019)
Adj. R^2	0.233	0.234	0.225	0.227	0.222	0.224	0.267	0.267
Clusters	3,539	3,539	3,539	3,539	3,539	3,539	3,539	3,539
N	164,912	164,912	164,912	164,912	164,912	164,912	164,912	164,912

Notes: **: $p < 0.01$, *: $p < 0.05$, ’: $p < 0.1$ (based on two-sided testing). The dependent variable is the external cost of transport aggregated to the person-day level. Standard errors (in parentheses) are clustered at the participant level. Precipitation is measured in mm per hour; heat and cold are as defined be eqs. (2)-(3). All regressions include fixed effects on the person, calendar day and day of study level.

Figure 5 displays the ATE in proportional (rather than absolute) terms for the different sources of external costs. The overall effect ("Pricing & Info") is highly statistically significant. The figure displays 80% confidence intervals, such that the probability that the true value lies to the right of the upper limit is 10%. Based on such one-sided testing, the effects of information alone ("Information") and of adding a price to information ("Difference") are also statistically significant at $p < 0.1$. This implies that both information and monetary incentives contribute to the overall effect.

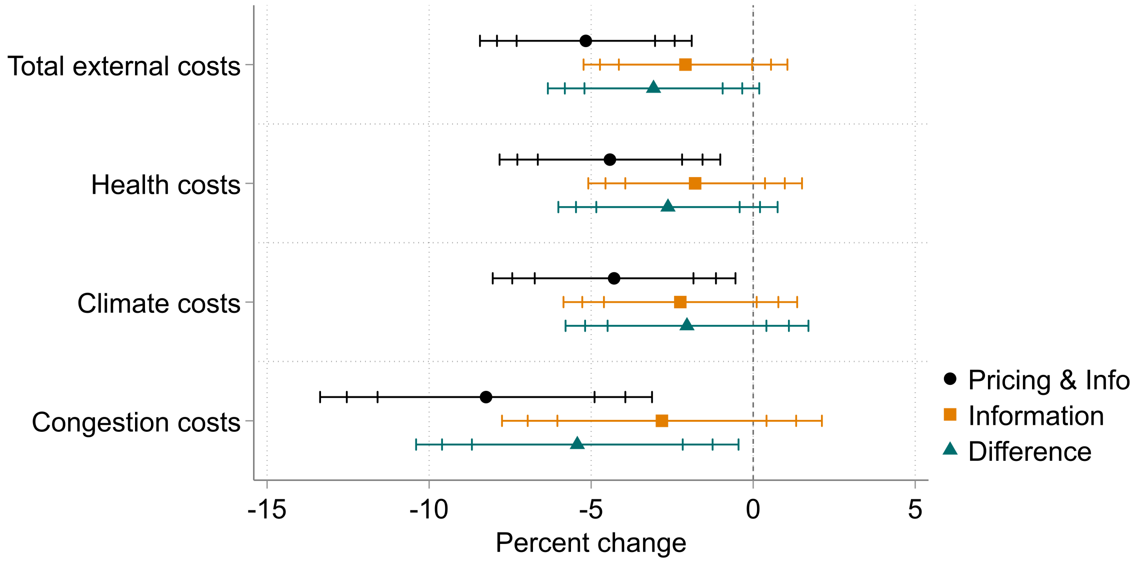
The proportional reduction of external health and climate costs is comparable to the reduction in total external costs, while the effect for congestion is somewhat stronger. As we will see below, this likely reflects the fact that congestion can be reduced not only by a mode shift, but also a shift in departure time.

The effect of information alone is not statistically significant for the sub-categories of external costs, presumably due to a lack of power. The effect of adding a price to information is weakly statistically significant (at $p<0.1$ for a one-sided test) except for climate externalities, and strongest for congestion. A possible explanation for this result is that reducing the external costs related to climate could be viewed as "the right thing to do" from a moral perspective, whereas the reducing the external costs of congestion simply means that someone else spends less time in traffic. This may not be as desirable from an altruistic point of view, such that the "pure money" dimension of the intervention carries relatively more weight. We will explore the heterogeneity of the effect below, including the role of personal "values".

Figure A.3 in the Appendix shows the treatment effect for the external costs associated with driving and public transport, and Table A.2 contains the underlying regression estimates. Whereas the proportional reduction of the external costs is similar for driving as for overall travel, we find no significant effect on the total external costs of PT travel, but a positive effect for health and climate costs. This suggests a mode shift that we will address in more detail below.

To interpret the magnitude of the ATE, we can compare this response to the change in the total price due to the implementation of the pricing scheme, i.e., the price including the private costs. We make the following assumptions: For public transport, we use the ticket price as a reference, either full or half fare depending on whether a particular participant holds a half-fare discount card. For participants that hold a public transport pass, we approximated the average cost by applying a discount to the half-fare ticket price. The level of the discount is determined by comparing the cost of a regional PT subscription

Figure 5: Treatment effect on the external costs of transport



Notes: The figure shows the average treatment effect for overall travel. The proportional effects were computed by scaling the regression coefficients in Table 5 by the average external cost of the control group during the treatment period. The bars show 80%, 90% and 95% confidence intervals, respectively.

with the corresponding cost if one were to buy a daily pass on 22 days per month.[16] For driving, we use an average value of CHF 0.70 per km, which is the official value used for deducting commuting expenses from taxable income and also very close to the value that the respondents in the final survey reported as their average costs of driving. We abstract from the purchase or rental price of bicycles and set the private monetary cost of active transport (cycling and walking) to zero.

Given these assumptions, the average daily private cost of transport for the control group during the treatment period was CHF 25.71 (see Table 4). The external cost was CHF 4.24, which corresponds to a price increase of 16.5%. Dividing the ATE by the average external costs leads to a proportional reduction of 5.2% (this is the point estimate in Figure 5). The resulting elasticity, in terms of a percentage reduction in external costs in response to a one-percent increase in the costs of transport is therefore -5.2/16.5=-0.32.[17] In words, this implies that introducing a transport pricing scheme

[16]The savings implicit in the subscription ranges from 24% in Geneva to 76% in Basel.

[17]This is the elasticity of *external costs* with respect to the *total price* of transport, which combines two different concepts. Note, however, that the elasticity of the external costs with respect to the externality-related price cannot be computed as the latter is zero before the treatment.

based on external costs that raises total transport costs, on average, by 10% would lead to a reduction in the external costs of transport by 3.2% in the short run.

Table 6 shows the sensitivity of the results with respect to controlling for the different fixed effects, the weather and using a control group. The preferred model is in column (1), followed by the model that adds weather controls in column (2). Removing either the day of study fixed effects (columns 3-4) or the calendar day fixed effects (columns 5-6) does not significantly change the results either; however, when removing both, the elasticity more than doubles (columns 7-8). Controlling for unobserved characteristics that vary over time is therefore crucial for identification. When estimating the ATE using only before-vs-after data for the pricing group, the resulting elasticity is around -0.37 if calendar day FE are included, and between -0.56 and -0.71 without (day of study FE cannot be included in this setting as they would be collinear with the "Post" dummy). This highlights the importance of including a control group in the experiment that is exposed to the same shocks as the treatment group. The elasticity is significantly over-estimated in the before-vs.-after setting, because the treatment also absorbs a part of the seasonal and day-of-study effects. Controlling for the weather or calendar day FE mitigates the problem, but it cannot fully remove the bias.

Table 6: Sensitivity analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Pricing	-0.220**	-0.222**	-0.226**	-0.230**	-0.221**	-0.217**	-0.464**	-0.536**				
	(0.071)	(0.070)	(0.063)	(0.063)	(0.071)	(0.071)	(0.052)	(0.053)				
Information	-0.089	-0.094	-0.095	-0.101'	-0.090	-0.087	-0.334**	-0.407**				
	(0.068)	(0.068)	(0.059)	(0.059)	(0.068)	(0.068)	(0.048)	(0.050)				
Precipitation		0.002		0.002		-0.003		-0.005*		0.003		-0.007'
		(0.004)		(0.004)		(0.002)		(0.002)		(0.007)		(0.004)
Heat		0.191**		0.191**		0.045**		0.027**		0.200**		0.047**
		(0.019)		(0.019)		(0.006)		(0.005)		(0.031)		(0.010)
Cold		-0.512**		-0.511**		-0.240**		-0.179**		-0.420**		-0.263**
		(0.074)		(0.074)		(0.038)		(0.037)		(0.135)		(0.073)
Post									-0.302**	-0.306**	-0.464**	-0.587**
									(0.105)	(0.104)	(0.052)	(0.057)
Prop. effect	-0.052**	-0.052**	-0.053**	-0.054**	-0.052**	-0.051**	-0.110**	-0.127**	-0.064**	-0.065**	-0.099**	-0.125**
	(0.017)	(0.017)	(0.015)	(0.015)	(0.017)	(0.017)	(0.013)	(0.013)	(0.022)	(0.022)	(0.010)	(0.011)
Elasticity	-0.315**	-0.318**	-0.323**	-0.329**	-0.317**	-0.311**	-0.665**	-0.768**	-0.366**	-0.371**	-0.563**	-0.711**
	(0.105)	(0.105)	(0.093)	(0.093)	(0.105)	(0.105)	(0.077)	(0.079)	(0.125)	(0.125)	(0.059)	(0.065)
Person FE	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑
Cal. day FE	☑	☑	☑	☑	☐	☐	☐	☐	☑	☑	☐	☐
Study day FE	☑	☑	☐	☐	☑	☑	☐	☐	☐	☐	☐	☐
Adj. R^2	0.233	0.234	0.232	0.234	0.229	0.230	0.229	0.229	0.228	0.229	0.224	0.225
Clusters	3,539	3,539	3,539	3,539	3,539	3,539	3,539	3,539	1,152	1,152	1,152	1,152
N	164,912	164,912	164,912	164,912	164,913	164,913	164,913	164,913	53,367	53,367	53,369	53,369

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors in parentheses and clustered at participant level. The dummy variable "post" takes the value of one during the treatment period (study days 29-56), and zero otherwise. In columns 1-8, the proportional values and elasticities are derived using the external costs and the price increases of the control group (CHF 4.24 and 16.4%, respectively). For the before-vs.-after analysis in columns (9)-(12), the daily external cost (CHF 4.51) and price increase (17.5%) was computed for the pricing group during the observation phase.

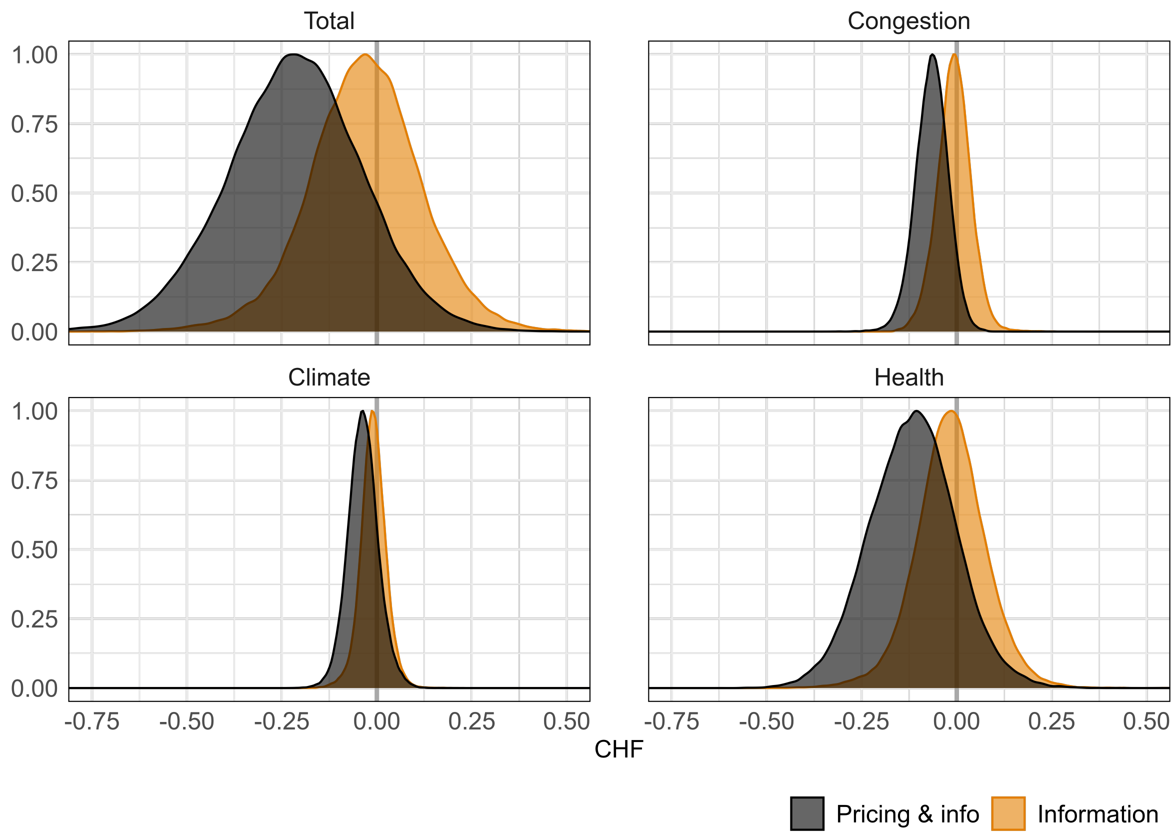
5.3 Effect heterogeneity

The overall ATE could mask heterogeneous responses within different segments of the population. In order to investigate a potential effect heterogeneity, we employ a "causal forest" approach based on the generalized regression forest algorithm proposed by Wager and Athey (2018) and implemented in the R package *grf* (Tibshirani, Athey and Wager, 2020). In contrast to the regression approach, the causal forest is agnostic as to which individual characteristics may generate heterogeneous treatment effects. The regression trees in the causal forest algorithm are grown by conditioning on those variables that generate the most heterogeneous treatment effects at each node. This procedure is repeated many times on samples randomly drawn without replacement from the data to form a causal forest. The average treatment effect is estimated by substituting the conditional predictions from the causal forest into the doubly robust augmented inverse probability weighting estimator proposed by Robins, Rotnitzky and Zhao (1994). The splits can be tallied across trees to arrive at a measure for the most *important* splitting variables, weighted by the level at which the splits occur. The earlier the split, the higher the weight assigned to that variable in the importance measure. This results in a list of "important" variables in the sense that they generate the strongest heterogeneity in the ATE.

Figure 6 shows the distribution of the conditional treatment effect using the causal forest algorithm, both for the pricing and the information-only treatment. The relative variable importance derived from the algorithm is shown in Figure A.4. To interpret this measure, we also included a continuous and a discrete random variable. The variables with a higher importance ranking than these random variables can be treated as likely candidates to explain the effect heterogeneity, since they contain "better than random noise" information.

Besides the socio-demographic variables collected in the introduction survey, we also included a number of variables from the final survey, which was conducted after the experiment was concluded. A battery of questions was used to elicit respondents' personal values (Schwartz, 1992; De Groot and Steg, 2010). Using this methodology, respondents were assigned an index along four dimensions labeled "altruistic", "egoistic", "hedonic" and "biospheric". Furthermore, we examined the extent to which the participants of the two treatment groups understood the concept of the external costs of transport that we explained to them at the beginning of the treatment, and in each of the reports. Specifically, we asked them to choose the definition of the external costs of transport from four possible answers. About 45% identified the correct answer, whereas the remainder

Figure 6: Distribution of conditional effects



Notes: The figure shows the distribution of the conditional treatment effects resulting from the causal forest approach for total external costs (top left) and the sub-categories considered.

answered incorrectly.[18]. We include this information as a dummy labeled "Correct EC".

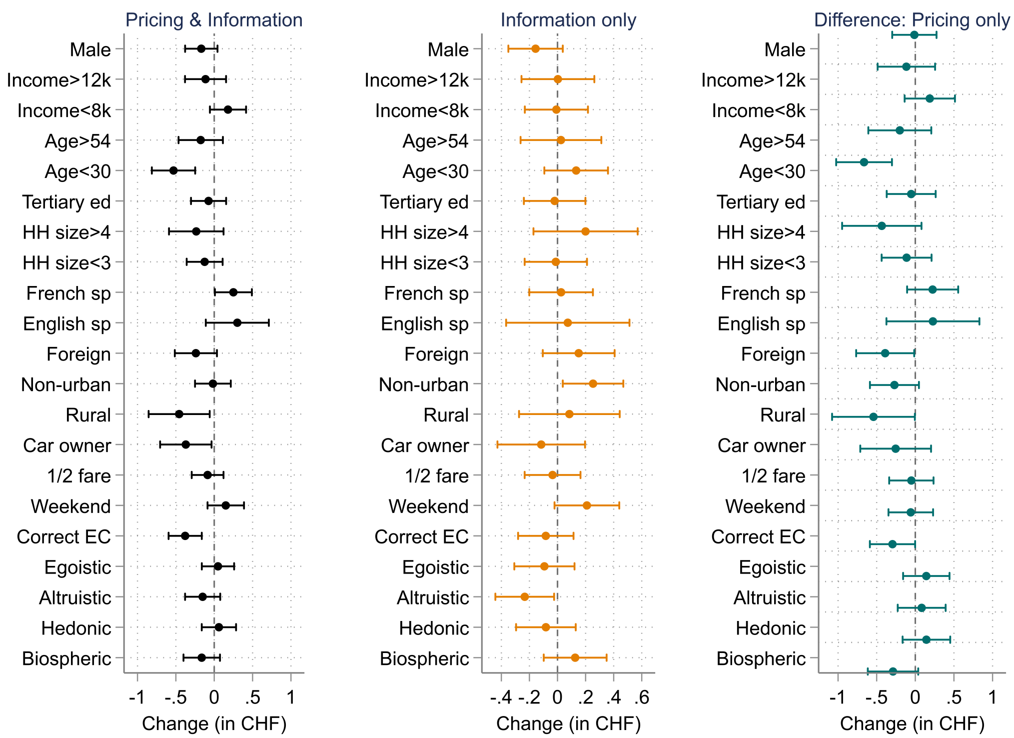
We include the variables identified to be "important" by the CF algorithm as interaction terms with the "Pricing" $\times$ "post" and the "Information" $\times$ "post" dummies. These variables are dummies based on gender, income, age, education, household size, language, citizenship, urbanization, car ownership, owning a half-fare public transport subscription, weekend, external cost question and the four values dimensions. To account for the correlation among these dummies, we include them jointly in a multi-variate regression.

Figure 7 displays the coefficients on the interaction terms, separately for the pricing

[18]The question was formulated as "How would you define the external costs of your transport behavior?" The possible responses were: (i) the costs associated with my travel behavior that I have to pay myself; (ii) the costs imposed on society as a result of my travel behavior; (iii) the total costs associated with my travel behavior (sum of private and societal costs); (iv) I don't know what the external costs of travel are

and information, information only and their difference; the underlying regression coefficients are in Table A.3. Overall, we find that the effect is relatively homogeneous across socio-demographic characteristics, with some exceptions. Setting p$\leq$ 0.05 as the threshold (two-sided testing), we find that the overall effect is stronger (i.e., more negative) for the young, those living in rural areas, car owners and those who correctly identified the definition of external transport costs. For French speakers, the effect is weaker.

Figure 7: Treatment effect heterogeneity



Notes: Results from including interaction dummies in a multivariate regression. The bars denote 95%-confidence intervals. The results are given in CHF per day. The underlying regression coefficients are shown in Table A.3.

The stronger overall response for the young, those living in rural areas and the respondents that answered the "exam" question correctly is due to a stronger response to pricing. For car owners, both effects seem to have contributed to the stronger effect, although neither of them is statistically significant by itself. This may simply be a sign that car owners have a greater potential to reduce external costs. For people that live in municipalities that have an intermediate level of urbanization ("Non-urban"), a higher responsiveness to information is neutralized by a lower responsiveness to pricing, such that

Table 7: Elasticities for subsamples

	Treatment effect (%)			Price increase (%)			Elasticity				
	Estimate	Lower Bound	Upper Bound	Estimate	Lower Bound	Upper Bound	Estimate	Lower Bound	Upper Bound	p	N
Age>54	-6.28	-14.02	1.45	16.22	15.83	16.62	-0.39	-0.87	0.09	0.112	32,221
30≤Age≤54	-2.92	-7.22	1.39	16.61	16.37	16.86	-0.18	-0.44	0.08	0.185	89,863
Age<30	-9.51	-16.32	-2.70	16.39	16.01	16.77	-0.58	-1.00	-0.17	0.006	42,822
German sp.	-7.34	-11.33	-3.35	16.31	16.10	16.52	-0.45	-0.69	-0.21	0.000	110,146
French sp.	1.31	-5.41	8.04	16.85	16.46	17.24	0.08	-0.32	0.48	0.702	42,666
English sp.	-5.38	-17.63	6.88	17.22	16.34	18.10	-0.31	-1.03	0.40	0.391	12,093
Urban	-3.53	-7.85	0.80	16.80	16.55	17.05	-0.21	-0.47	0.05	0.110	106,009
Non-urban	-5.98	-11.96	0.01	16.14	15.81	16.47	-0.37	-0.74	0.00	0.051	45,464
Rural	-12.42	-22.40	-2.44	15.71	15.19	16.23	-0.79	-1.43	-0.15	0.015	13,426
Car owner	-5.06	-8.31	-1.80	16.46	16.27	16.66	-0.31	-0.50	-0.11	0.002	144,680
No car owner	-6.34	-19.12	6.44	16.68	16.04	17.33	-0.38	-1.15	0.39	0.331	20,231
Incorrect EC	-0.35	-3.97	3.26	16.48	16.29	16.67	-0.02	-0.24	0.18	0.848	117,329
Correct EC	-10.30	-14.73	-5.87	16.48	16.29	16.67	-0.63	-0.89	-0.36	<0.001	105,892

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors in parentheses and clustered at participant level. The dummy variable "post" takes the value of one during the treatment period (study days 29-56), and zero otherwise.

the overall effect is almost exactly the same as that of the reference category "urban".

Last, we find that the study participants that scored above the median in terms of the altruistic index responded significantly more to information alone, which is consistent with expectations. We see no differential response of altruists to pricing (suggesting that there is no"crowding out" effect). There were no statistically significant differences along the other three values-dimensions.

Table 7 presents the proportional effects, price increases and resulting elasticities for the sub-samples for which we found statistically significant differences in terms of the overall effect (pricing & information). The elasticity of the participants that correctly identified the external costs in the "exam" question is -0.64, whereas the elasticity of the rest of the sample (around 55% of the participants in the treatment groups) is precisely centered around zero. This implies that the ATE is exclusively driven by those participants that understood the nature of the experiment. The table also shows the elasticity for different subsamples defined by age, language, car ownership and the level of urbanity or a respondents' home municipality.

5.4 Mechanisms

People can reduce their external costs of transport in different ways: Travel less, substitute towards modes associated with lower external costs and choose different routes and departure times. To shed light on potential mechanisms, Figure **??** shows the effect of the pricing treatment on various outcomes of interest. The regression results underlying this figure are shown in Tables A.4 - A.7 in the Appendix.

The treatment did not significantly reduce overall travel distances, but we measure a statistically significant reduction in car distance countered by increases in the other modes (panel a). The effect can be seen separately on the intensive margin in panel b (i.e., conditional on traveling with a particular mode on a given day) and on the extensive margin in panel c (i.e., the probability of traveling). The mode shift becomes more salient if the treatment effect is shown for mode share in terms of distance (panel d). There is a statistically significant reduction in the share of car distance by about 3% and an increase in the share of public transport, bicycling and walking.

The pricing treatment significantly reduced congestion costs per km of car travel (panel e), implying that modal shift is not the only mechanism responsible for the reduction in external costs. The reduction in congestion per km can be due to a change in route and/or a change in departure time. Using the departure time (in minutes) as the dependent variable, we observe a significant shift in the departure times for car trips in the morning towards earlier departures, but no effect in the evening (panel f).[19] There was no reduction in crowding for public transport, nor was there an effect on PT departure time.

5.5 Welfare effects

The reduction in external costs due to treatment is a welfare gain. However, people experience disutility from changing their behavior in response to the pricing. We therefore have to compute the change in private utility in response to the pricing and then subtract this from the gain via external costs.

To do this, we estimate a mode choice regression for each trip in our database. As a first step, this requires the generation of non-chosen alternatives. We have completed this step and are now in the process of estimating mixed logit models. Once these models are estimated, we can derive the change in utility as the change in the log sums from the model, as in (Kreindler, 2023)

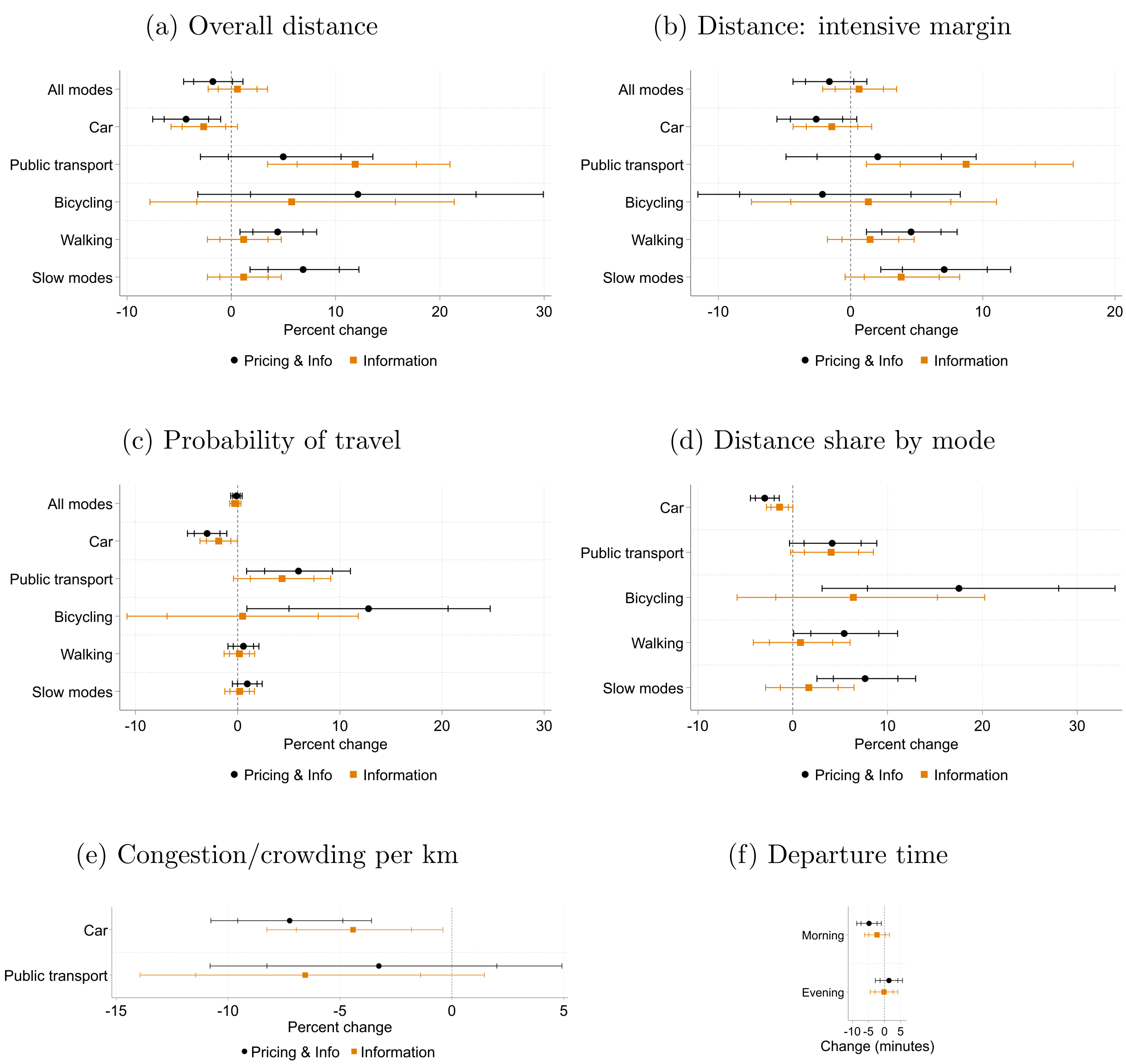
This part of the paper is work in progress.

5.6 Acceptability

Even if transport pricing works, its implementation may be challenging not only in terms of technology and data confidentiality, but also in terms of social acceptability. In the

[19]For this regression, we only included people for whom we observed at least one peak-hour trip during the observation phase. We then computed the average trip departure time for trips before and after noon.

Figure 8: Mechanisms underlying the reduction in external costs



Notes: The bars denote 80%- and 95%-confidence intervals. In panels (a), (b), (d) and (e), the treatment effects are computed using a Poisson pseudo-maximum likelihood (ppml) regression. Panel (c) shows the marginal results (semi-elasticity) of a logit regression. In panel (f), a linear DiD-specification is chosen with the departure time (measured in minutes after midnight) as the dependent variable. Slow modes is the sum of cycling and walking.

introduction survey, we included three questions designed to elicit respondents' preferences for a possible introduction of transport pricing. In order not to reveal the purpose of the experiment, these questions were part of a larger number of queries posed to the respondents. The three questions presented the same concept but were worded differently:

Please indicate whether you agree or disagree with each policy:

Time- and route-specific mobility pricing, made revenue-neutral by lowering other taxes.

Please indicate your level of agreement or disagreement with the following statements:

The price for mobility should reflect the social cost (e.g., health, environment, congestion).
The transport network should be used more efficiently by introducing dynamic pricing.

Figure 9 shows the responses. A majority of the respondents were either positive or neutral if the question was worded with respect to social costs or revenue-neutral transport (or mobility) pricing.[20] However, if the focus was placed on the time-varying nature of this pricing, the majority were opposed. This suggests that transport pricing could, in principle, find a political majority, but that it depends on how it is communicated.

The majority of the participants preferred at least some of the money to be used for transport projects, whereas fewer than 8% indicated that they preferred the revenue to be returned to households (Fig. A.2). The revenue neutrality embedded in the question about mobility pricing thus likely led to a lower level of approval, *ceteris paribus*.

6 Challenges to identification and external validity

In this section, we address problems that could potentially threaten the validity of our results internally and externally.

6.1 Observation effect

As mentioned in Section 5.1, we observe a reduction in the external costs of travel during the course of the experiment. This is partly due to a seasonal effect, but the effect persists even after controlling for seasonality. Table 8 shows the results from regressing the daily external costs of travel during the observation phase on dummies denoting membership in

[20]The same questions were asked again in the final survey, which took place after the experiment. Despite the different samples (21,800 respondents of the introduction survey vs. 3,521 participants in the tracking study), the answers were qualitatively similar. There was also no systematic difference in the answers of the three experimental groups.

Figure 9: Support for transport pricing

Notes: The figure shows participants' responses to the questions described in the main text.

the Pricing and the Information groups, calendar day FE and a linear day-of-study trend. This trend is significant and indicates that the external costs from all travel decrease by 1.5 cents per study day (column 1). However, there is no difference in this trend across the groups (column 2), which is crucial for identification purposes. Columns 3-6 show that the trend is caused by a reduction in the external costs of driving. When using distance as the dependent variable (Table A.8), the results indicate that there may be a shift away from driving and towards public transport during the pre-treatment period, but this shift is the same for all groups.

6.2 Strategic mode corrections

One possible explanation for the trend discussed in the previous section is a strategic correction of modes. Participants were invited to use the validation interface to confirm the detected mode and purpose of their trips and activities. As the mode is crucial in determining the external costs, the possibility of overwriting the detected mode provided an opportunity for the participants in the pricing group to "game" the experiment, e.g., by mis-assigning actual car trips to another transport mode.

On the other hand, mode adjustments could also be true corrections of a wrongly detected mode. The key question is therefore whether we observe systematically different

Table 8: Trends in external costs

	All Travel	All Travel	Car	PT	Bicycle	Walking
Info	9.200	-0.291	-3.196	4.485	-0.346	-1.234’
	(12.179)	(15.935)	(16.143)	(2.849)	(0.820)	(0.682)
Pricing	19.573	7.062	2.182	5.830’	0.247	-1.197’
	(12.344)	(15.968)	(16.154)	(2.988)	(0.947)	(0.661)
Day of study trend	-1.177**	-1.663**	-1.671**	0.097	-0.007	-0.081**
	(0.395)	(0.535)	(0.538)	(0.102)	(0.027)	(0.024)
Pricing x Day-of-study trend		0.836	0.764	0.053	-0.022	0.041
		(0.633)	(0.638)	(0.101)	(0.032)	(0.029)
Info x Day-of-study trend		0.634	0.714	-0.111	-0.004	0.036
		(0.638)	(0.641)	(0.102)	(0.031)	(0.029)
Adj. R^2	0.005	0.005	0.005	0.009	0.007	0.014
N	83,855	83,855	83,855	83,855	83,855	83,855

Notes: **: $p < 0.01$, *: $p < 0.05$, ’: $p < 0.1$. The dependent variable is the external cost of transport (in cents per day), aggregated to the person-day level, and restricted to study days 1-28 (pre-treatment). Standard errors (in parentheses) are clustered at the participant level. The regressions additionally include calendar day FE.

mode correction behavior for the pricing group relative to the control and information groups. To test for this, we regress the number of daily mode corrections during the treatment phase on dummies indicating membership in the pricing and information groups, while controlling for day of sample and day of calendar day FE. We focus on the treatment phase because there was no incentive to act strategically during the observation period. The results are shown in column 1 of Table 9. We see no difference in the number of corrections per day across groups, and the results remain stable if we add a series of control variables (column 2).[21] Columns 3-4 show the marginal effects of a logit regression, using the same explanatory variable but with a dummy that is equal to one if a person has corrected a mode on a given day and zero otherwise. We find no differential correction behavior on the extensive margin.

To further test the robustness of our results, we rerun our base regression after removing all observations (on the person-day-level) that contain at least one mode correction (this removes 9.4% of the data in phase 2). The resulting treatment effects, proportional effects and elasticities are shown in column 2 of Table 10, along with the baseline results. Although the point estimates change somewhat for the different samples, the effects remain largely stable. This implies that our results are unlikely to be driven by strategic mode correction.

[21] These are the variables for which we compute the interaction terms in section 5.3.

Table 9: Mode Correction

	Corrections	Corrections	Probability (Corr.)	Probability (Corr.)
Pricing	1.027	1.009	0.002	0.000
	(0.068)	(0.066)	(0.006)	(0.006)
Information	0.984	0.982	0.000	-0.001
	(0.064)	(0.064)	(0.006)	(0.006)
Controls	☐	☑	☐	☑
N	77,704	75,511	77,662	75,469

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors (in parentheses) clustered at participant level. The dependent variable in columns 1-2 is the number of mode corrections per day. The coefficients are proportional effects, estimated using a ppml model (with 1.0 representing no effect). Columns 3-4 display the marginal effects from logit regressions. All regressions control for person, calendar and study day FE. The controls are those shown in Fig. 7.

Table 10: Subsample analyses

	Baseline	w/o corrections	w/o weeks 7-8	w/o weeks 5-6	w/o zeroes
Pricing	-0.220**	-0.231**	-0.243**	-0.230**	-0.217**
	(0.071)	(0.079)	(0.082)	(0.084)	(0.073)
Information	-0.089	-0.067	-0.088	-0.086	-0.088
	(0.068)	(0.077)	(0.080)	(0.082)	(0.070)
Constant	4.483**	4.725**	4.549**	4.504**	4.672**
	(0.019)	(0.020)	(0.015)	(0.016)	(0.020)
Prop. effect	-0.051	-0.054	-0.055	-0.049	-0.053
	(0.017)	(0.018)	(0.019)	(0.021)	(0.017)
Elasticity	-0.310	-0.329	-0.339	-0.300	-0.324
	(0.103)	(0.109)	(0.116)	(0.126)	(0.103)
Clusters	3.539	3.539	3.539	3.539	3.539
Adj. R^2	0.233	0.238	0.236	0.232	0.238
N	164,912	143,333	123,485	125,281	158,116

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors in parentheses and clustered at participant level. All regressions include fixed effects for person, day of study and day of calendar. The proportional effect and the elasticity are computed using the averages of the control group subject to the appropriate restrictions.

6.3 Self-selection due to nonrandom attrition

The assignment into groups was randomized, but people chose whether or not they wanted to continue with the study. If people in the pricing group were more or less likely to drop out of the study than the control group, this could lead to a bias due to self-selection. We address this issue in detail in Appendix B.8, where we investigate participant retention and find no differential attrition across groups.

Furthermore, if attrition were influenced by the group assignment, any bias due to self-selection should increase over the course of the treatment phase. We therefore re-estimate our base model using only the first two weeks (column 3 of Table 10) or the last two weeks of the treatment period (column 4). The results remain largely unchanged, indicating that nonrandom attrition is unlikely to be an important source of bias.

6.4 Treatment of missing tracking data

Many participants did not deliver tracks on all days. To differentiate between true zeroes (i.e., participants staying at home) and missings (participants disabling the app), we rely on imputed activities. Suppose that a participant travels home on Friday evening and does not deliver another track until Monday. If the app imputes an uninterrupted activity "at home" that lasts from Friday to Monday, then we assume that this person stayed at home and assign a travel distance of zero. However, the imputation of activities and locations is not always correct. To gauge the sensitivity or our results to the issue of missings vs. zeroes, we re-estimate the model using only data from days with nonzero travel distances. The resulting ATE is shown in column 5 of Table 10. As it this very similar to the baseline, the ATE is not driven by missing data mistakenly coded as zero.

6.5 External validity

Every study is externally valid for some setting and no study is externally valid in all settings (List, 2020). For a study to provide useful insights beyond its immediate setting, List argues that the burden of proof for authors of empirical work consists of four transparency conditions: (1) selection, (2) attrition, (3) naturalness and (4) scaling.

We find no differential attrition between the groups (see above), and since our experiment did not introduce new tasks but simply observed people in their everyday travel, the naturalness condition is arguably not a problem here. In the following, we will therefore focus on the selection and scaling conditions.

Selection

Our sample does not significantly differ from the general population living in Swiss urban areas in terms of socio-demographic characteristics (see section 2.1). However, due to self-selection into the tracking-part of the study, it is possible that it differs from the target population in terms of unobservable characteristics such as constraints, beliefs and preferences, all of which determine choices. For this reason, we cannot guarantee external validity given our sample selection procedure.

However, we stress that we were careful not to make any reference to transport pricing or external costs when inviting people to participate in the tracking study. In order for our results to mis-represent the response of the target population, there would need to be a correlation between the propensity to participate in a tracking study and the extent to which someone responds to information and pricing associated with the external costs of transport. Although this is possible, the fact that our treatment effects are homogeneous across most socio-economic characteristics suggests that this is unlikely to be a large source of bias. One exception is the share of young adults (ages 18-25), which is somewhat higher in our sample. Since people below age 30 tend to respond more to pricing (see Table A.3), this could over-estimate the treatment effect relative to the effect that would be expected in the overall population.[22]

Last, the fact that tracking a person's transport choices requires consent (and thus self-selection) may also affect an eventual implementation. For example, if faced with political opposition due to privacy concerns, transport pricing could be offered to volunteers who, in exchange, are exempt from vehicle registration taxes (or receive some other lump-sum compensation). In such an implementation, the target population could be quite similar to our sample, such that self-selection in our study would become a feature rather than a source of bias.

Scaling

The scaling condition is related to the selection in our context, but it additionally raises the question about the effect of transport pricing for people that (i) do not drive regularly, (ii) do not live in urban areas, (iii) are outside the 18-65 age range and (iv) reside in different countries (and thus may have access to fewer public transport options than our Swiss sample).

[22]Note, however, that the over-sampling only pertains to the youngest age group; when combining the age categories 18-25 and 25-35, our sample is similar to the MTMC and thus representative of the target population.

Drivers were over-represented in this study by design. As driving produces the largest external costs, the observed short-term elasticity therefore may lead to over-estimating pricing effects in the overall population. On the other hand, congestion provides an important justification for transport pricing as opposed to a more simple fuel or per-km surcharge. Since congestion is mainly a problem of working-age drivers in urban areas, knowing the effect of transport pricing in a population living and driving in these areas is a useful starting point. Furthermore, the fact that car owners and residents in rural municipalities responded more strongly to the pricing treatment suggests that transport pricing reduces the production of external transport costs also in other settings. In order to derive the exact magnitude of the response, however, further studies are needed to replicate and extend our findings to other populations.

Since the pricing scheme in the experiment consisted of taking money away from a given budget, loss aversion may have increased the effect relative to a tax (Tversky and Kahneman, 1991). On the other hand, there is some evidence that people treat "house (gambling) money" differently from "real" money. Thaler and Johnson (1990) show that individuals tend to integrate (combine) prior gains (in this study the accumulated budget over the 4 week observation phase) with subsequent losses, which, as long as they are less than the initial gain, are seen as "reductions in a gain" rather than a loss. This facilitates risk-seeking behavior until the prior gain is completely depleted. For our study, this would suggest risk-seeking behavior among participants in the pricing group, since the budget was "house money" until we paid participants the balance in "real" money at the conclusion of the study. This would also lead to an under-estimate of the effect relative to transport pricing that would become part of households' general expenditure.

Most importantly, the treatment in our experiment lasted only one month, and there are a number of arguments for expecting larger effects in the long run. With a permanent introduction of transport pricing, additional margins of response will become available such as the choice of work and home locations, changes in activity routines, vehicle/transit pass ownership or negotiations with employers about work hours and location. One would expect potential welfare-improving general equilibrium effects such as increased car travel speeds and less crowding in public transport during peak hours to materialize as transport pricing is scaled to a larger portion of the population. Furthermore, the behavioral response was concentrated among those respondents who understood the concept of external costs underlying the pricing. Whereas it is to be expected that not everyone pays close attention to the "rules" in a short study, a general introduction of transport pricing would presumably have a greater salience.

For policymakers, other avenues of revenue generation for transport infrastructure are gaining importance as the share of electric vehicles increases and revenue from fuel taxes and surcharges decreases.[23] Transport pricing on a larger scale may alleviate these concerns since congestion, noise, and emissions of local pollutants (through braking and tire wear) are external costs of car travel, regardless of fuel type.

7 Conclusion

The MOBIS experiment implemented transport pricing based on the social marginal costs. The external health costs were the most important, followed by congestion and climate costs. The short-term elasticity for total external costs associated with the pricing treatment was -0.32. Whereas the information-only treatment had a strong effect for subgroups of the population (such as altruists), the effect is only marginally statistically significant for the sample overall. However, our results imply that both information and monetary incentives play an important role in explaining the behavioral change in our experiment. The reduction in the external costs is due to a combination of a shift away from driving towards other modes and towards less congested times and routes. The effect varies with age, degree of urbanity, car ownership and language region, and particularly strongly with the degree to which participants engaged with the experiment. Those that understood the concept of external costs, and thus the pricing mechanism, responded twice as strongly relative to those that did not.

The elasticity estimate is comparable to results based on toll pricing (Bain, 2019), but lower than earlier estimates based on before-vs.-after studies (Leape, 2006; Nielsen, 2004). The MOBIS experiment is the first multi-modal RCT investigating pricing in a transport setting and thus different to uni-modal pricing schemes, where the lack of pricing for alternative modes may have inflated the mode shift effect of the pricing scheme.

Our experiment shows that multi-modal transport pricing works in practice. The required technology is available, and a number of countries have computed the external costs of mobility within their borders. The COVID-19 pandemic has demonstrated that patterns of living, working and traveling are more adjustable than previously assumed. It seems justified to expect people to respond to the price incentives in similar, albeit less dramatic ways. Furthermore, a transition away from the current transport funding that relies mostly on fuel taxes is unavoidable due to shifts in modes, fuel types and vehicle technologies. Pigovian transport pricing is an alternative funding mechanism that can also be implemented in the presence of a sizeable electric vehicle fleet.

[23] The fuel tax and fuel surcharge generated 71% of total revenue for the Federal Roads Office in 2021 (Federal Roads Office - ASTRA, 2021).

A Pigovian pricing scheme as used in the MOBIS experiment would face a number of challenges for practical implementation due to privacy concerns, limited social acceptability and the technical constraints of assessing the tax on a real-time basis (including an update of the congestion costs, which will change if pricing leads to significant peak shifting). However, even a simplified pricing scheme should be guided by the marginal external costs of transport to increase the efficiency of the transport system. A key challenge will be to agree on the price setting (e.g., the value of time or the social cost of carbon) within the political process. Furthermore, it is well-known that fuel taxes are regressive (West and Williams, 2004; Bento et al., 2009), and the distributional aspects of a cost-based pricing scheme like the one used here thus deserve further investigation. Efforts to advance such a scheme will need to be complemented with re-distributive measures to counteract adverse distributional implications.

Multi-modal transport pricing based on the external costs of transport is feasible and has the desired effect of shifting modes, departure times and routes. It thus leads to a more efficient use of the transport system and a reduction in the need for network expansions. If implemented in an equitable way, transport pricing could become a key pillar of sustainable transport policy.

References

Agarwal, Sumit, and Kang Mo Koo. 2016. "Impact of electronic road pricing (ERP) changes on transport modal choice." *Regional Science and Urban Economics*, 60: 1–11.

Anderson, Michael L. 2014. "Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion." *American Economic Review*, 104(9): 2763–96.

Bain, Robert. 2019. "Toll Road Pricing: Demand Elasticity and Affordability A State-of-the-Practice Scan." Mimeo.

Beheshtian, Arash, R. Richard Geddes, Omid M. Rouhani, Kara M. Kockelman, Axel Ockenfels, Peter Cramton, and Wooseok Do. 2020. "Bringing the efficiency of electricity market mechanisms to multimodal mobility across congested transportation systems." *Transportation Research Part A: Policy and Practice*, 131: 58–69. Developments in Mobility as a Service (MaaS) and Intelligent Mobility.

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

Bento, Antonio, Daniel Kaffine, Kevin Roth, and Matthew Zaragoza-Watkins. 2014. "The effects of regulation in the presence of multiple unpriced externalities: Evidence from the transportation sector." *American Economic Journal: Economic Policy*, 6(3): 1–29.

Bento, Antonio M, Lawrence H Goulder, Mark R Jacobsen, and Roger H Von Haefen. 2009. "Distributional and efficiency impacts of increased US gasoline taxes." *American Economic Review*, 99(3): 667–99.

Börjesson, Maria, and Ida Kristoffersson. 2018. "The Swedish congestion charges: Ten years on." *Transportation Research Part A: Policy and Practice*, 107: 35–51.

Bösch, Patrick M, Kirill Müller, and Francesco Ciari. 2016. "The ivt 2015 baseline scenario." 16th Swiss Transport Research Conference (STRC 2016).

Bothos, Efthimios, Gregoris Mentzas, Sebastian Prost, Johann Schrammel, and Kathrin Röderer. 2014. "Watch your Emissions: Persuasive Strategies and Choice Architecture for Sustainable Decisions in Urban Mobility." *PsychNology Journal*, 12(3).

Carreras, Iacopo, Silvia Gabrielli, Daniele Miorandi, Andrei Tamilin, Fabio Cartolano, Michal Jakob, and Stefano Marzorati. 2012. "SUPERHUB: a user-centric perspective on sustainable urban mobility." 9–10.

CE Delft. 2019. "Handbook on the external costs of transport: Version 2019." European Commission. Publication 18.4K83.131.

Cellina, Francesca, Dominik Bucher, Francesca Mangili, José Veiga Simão, Roman Rudel, and Martin Raubal. 2019. "A large scale, app-based behaviour change experiment persuading sustainable mobility patterns: Methods, results and lessons learnt." *Sustainability*, 11(9): 2674.

Correia, Sergio, Paulo Guimarães, and Thomas Zylkin. 2019. "Verifying the existence of maximum likelihood estimates for generalized linear models." *arXiv:1903.01633*.

Correia, Sergio, Paulo Guimarães, and Tom Zylkin. 2020. "Fast Poisson estimation with high-dimensional fixed effects." *The Stata Journal*, 20(1): 95–115.

Creutzig, Felix, Patrick Jochem, Oreane Y Edelenbosch, Linus Mattauch, Detlef P van Vuuren, David McCollum, and Jan Minx. 2015. "Transport: A roadblock to climate change mitigation?" *Science*, 350(6263): 911–912.

Davis, Lucas W. 2017. "Saturday driving restrictions fail to improve air quality in Mexico City." *Scientific reports*, 7: 41652.

De Groot, Judith IM, and Linda Steg. 2010. "Relationships between value orientations, self-determined motivational types and pro-environmental behavioural intentions." *Journal of Environmental Psychology*, 30(4): 368–378.

Dixit, Vinayak V, Andreas Ortmann, E Elisabet Rutström, and Satish V Ukkusuri. 2017. "Experimental Economics and choice in transportation: Incentives and context." *Transportation Research Part C: Emerging Technologies*, 77: 161–184.

Duranton, Gilles, and Matthew A Turner. 2011. "The fundamental law of road congestion: Evidence from US cities." *American Economic Review*, 101(6): 2616–52.

EEA. 2019. "Emissions of Air Pollution from Transport." European Environment Agency, European Environment Agency. Indicator Assessment.

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

Eliasson, Jonas, Lars Hultkrantz, Lena Nerhagen, and Lena Smidfelt Rosqvist. 2009. "The Stockholm congestion–charging trial 2006: Overview of effects." *Transportation Research Part A: Policy and Practice*, 43(3): 240–250.

Federal Roads Office - ASTRA. 2017. "Handbook NISTRA 2017." Federal Roads Office - ASTRA Technical Report, Bern.

Federal Roads Office - ASTRA. 2021. "Financing of Roads." Federal Roads Office - ASTRA Technical Report, Bern.

Federal Statistical Office. 2017. "Raumgliederungen der Schweiz (Swiss Area Characteristics)." Swiss Federal Statistical Office. `https://www.atlas.bfs.admin.ch/maps/13/de/12362_12361_3191_227/20389.html`.

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

Goodwin, Phil, Joyce Dargay, and Mark Hanly. 2004. "Elasticities of road traffic and fuel consumption with respect to price and income: a review." *Transport reviews*, 24(3): 275–292.

Götschi, Thomas, and Beat Hintermann. 2014. "Valuing public investments to support bicycling." *Swiss Journal of Economics and Statistics*, 150(4): 297–329.

Götschi, Thomas, Jan Garrard, and Billie Giles-Corti. 2016. "Cycling as a part of daily life: a review of health perspectives." *Transport Reviews*, 36(1): 45–71.

Hülsmann, F, R Gerike, K Kickhöfer, and Raphael Luz. 2011. "Towards a multi-agent based modeling approach for air pollutants in urban regions." 144–166, FGSV Verlag GmbH.

IEA. 2020. "Tracking Transport 2020." International Energy Agency, International Energy Agency.

INRIX. 2020. "Global Traffic Scorecard 2019." Trevor Reed, INRIX Research.

Isaksen, Elisabeth T., and Bjorn G. Johansen. 2021. "Congestion pricing, air pollution, and individual-level behavioral responses." Working paper, available at SSRN 3832230.

Jariyasunant, Jerald, Maya Abou-Zeid, Andre Carrel, Venkatesan Ekambaram, David Gaker, Raja Sengupta, and Joan L Walker. 2015. "Quantified traveler: Travel feedback meets the cloud to change behavior." *Journal of Intelligent Transportation Systems*, 19(2): 109–124.

Kaddoura, Ihab. 2015. "Marginal congestion cost pricing in a multi-agent simulation investigation of the greater Berlin area." *Journal of Transport Economics and Policy*, 49(4): 560–578.

Karich, Peter, and Stefan Schröder. 2014. "Graphhopper." *http://www. graphhopper. com, last accessed*, 4(2): 15.

Kickhöfer, Benjamin, Friederike Hülsmann, Regine Gerike, and Kai Nagel. 2013. "Rising car user costs: comparing aggregated and geo-spatial impacts on travel demand and air pollutant emissions." In *Smart Transport Networks.* Edward Elgar Publishing.

Knight, Frank H. 1924. "Some fallacies in the interpretation of social cost." *The Quarterly Journal of Economics*, 38(4): 582–606.

Kreindler, Gabriel. 2023. "Peak-hour road congestion pricing: Experimental evidence and equilibrium implications." National Bureau of Economic Research.

Kreindler, Gabriel E. 2018. "The welfare effect of road congestion pricing: Experimental evidence and equilibrium implications." *Mimeo.*

Kristal, Ariella S, and Ashley V Whillans. 2020. "What we can learn from five naturalistic field experiments that failed to shift commuter behaviour." *Nature Human Behaviour*, 4(2): 169–176.

Leape, Jonathan. 2006. "The London congestion charge." *Journal of Economic Perspectives*, 20(4): 157–176.

List, John A. 2020. "Non est disputandum de generalizability? a glimpse into the external validity trial." National Bureau of Economic Research.

Maerivoet, Sven, Frank Daems, Friedl Maertens, Karel Renckens, Philip Van Houtte, and Louis Buelens. 2012. "A Field Trial on Smart Mobility." *Procedia - Social and Behavioral Sciences*, 54: 926–935.

Martin, Leslie A, and Sam Thornton. 2017. "Can Road Charges Alleviate Congestion?" *Available at SSRN 3055428.*

Molloy, Joseph, Christopher Tchervenkov, Beat Hintermann, and Kay W Axhausen. 2020. "Tracing the Sars-CoV-2 impact: The first month in Switzerland." *Transport Findings.*

Möser, Guido, and Sebastian Bamberg. 2008. "The effectiveness of soft transport policy measures: A critical assessment and meta-analysis of empirical evidence." *Journal of Environmental Psychology*, 28(1): 10–26.

Nielsen, Otto Anker. 2004. "Behavioral Responses to Road Pricing Schemes: Description of the Danish AKTA Experiment." *Journal of Intelligent Transportation Systems*, 8(4): 233–251.

Nikolic, Marija, and Michel Bierlaire. 2017. "Review of transportation mode detection approaches based on smartphone data."

Oum, Tae Hoon, William G Waters, and Jong-Say Yong. 1992. "Concepts of price elasticities of transport demand and recent empirical estimates: an interpretative survey." *Journal of Transport Economics and policy*, 139–154.

Parry, Ian WH, and Kenneth A Small. 2005. "Does Britain or the United States have the right gasoline tax?" *American Economic Review*, 95(4): 1276–1289.

Parry, Ian WH, Margaret Walls, and Winston Harrington. 2007. "Automobile externalities and policies." *Journal of economic literature*, 45(2): 373–399.

Pigou, Arthur Cecil. 1920. *The economics of welfare.* Macmillan & Co.

Pluntke, Christopher, and Balaji Prabhakar. 2013. "INSINC: a platform for managing peak demand in public transit." *JOURNEYS, Land Transport Authority Academy of Singapore*, 31–39.

Portney, Paul R, Ian WH Parry, Howard K Gruenspecht, and Winston Harrington. 2003. "Policy watch: the economics of fuel economy standards." *Journal of Economic perspectives*, 17(4): 203–217.

Robins, James M, Andrea Rotnitzky, and Lue Ping Zhao. 1994. "Estimation of regression coefficients when some regressors are not always observed." *Journal of the American statistical Association*, 89(427): 846–866.

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

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

Schwartz, Shalom H. 1992. "Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries." In *Advances in experimental social psychology.* Vol. 25, 1–65. Elsevier.

Small, K. A. 2008. "Urban transportation policy: A guide and road map." Wharton School, University of Pennsylvania.

Small, K. A., Erik Verhoef, and Robin Lindsey. 2007. *The economics of urban transportation.* Routledge.

Swiss Federal Office of Statistics and Swiss Federal Office of Spatial Development. 2017. "Mobility and Transport Microcensus 2015." `https://www.are.admin.ch/are/en/home/mobility/data/mtmc.html`.

Thaler, Richard H, and Eric J Johnson. 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.

Tibshirani, Julie, Susan Athey, and Stefan Wager. 2020. "grf: Generalized Random Forests." R package version 1.2.0.

Tirachini, Alejandro, and David A Hensher. 2012. "Multimodal transport pricing: first best, second best and extensions to non-motorized transport." *Transport Reviews*, 32(2): 181–202.

Tirachini, Alejandro, David A Hensher, and John M Rose. 2013. "Crowding in public transport systems: effects on users, operation and implications for the estimation of demand." *Transportation Research Part A: Policy and Practice*, 53: 36–52.

Tversky, Amos, and Daniel Kahneman. 1991. "Loss aversion in riskless choice: A reference-dependent model." *The Quarterly Journal of Economics*, 106(4): 1039–1061.

Van Benthem, Arthur. 2015. "What is the optimal speed limit on freeways?" *Journal of Public Economics*, 124: 44–62.

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

Vickrey, William S. 1963. "Pricing in urban and suburban transport." *The American Economic Review*, 53(2): 452–465.

Vrtic, Milenko, Nadine Schüssler, Alexander Erath, Michaela Bürgle, Kay W Axhausen, Emma Frejinger, Jelena Stojanovic, Michel Bierlaire, Roman Rudel, Stefano Scagnolari, et al. 2008. "Einbezug von Reisekosten bei der Modellierung des Mobilitätsverhaltens." *Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK, Bundesamt für Strassen*, 1191.

Wager, Stefan, and Susan Athey. 2018. "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests." *Journal of the American Statistical Association*, 113(523): 1228–1242.

West, Sarah E, and Roberton C III Williams. 2004. "Estimates from a consumer demand system: implications for the incidence of environmental taxes." *Journal of Environmental Economics and management*, 47(3): 535–558.

Wu, Linlin, Biao Yang, and Peng Jing. 2016. "Travel mode detection based on GPS raw data collected by smartphones: A systematic review of the existing methodologies." *Information*, 7(4): 67.

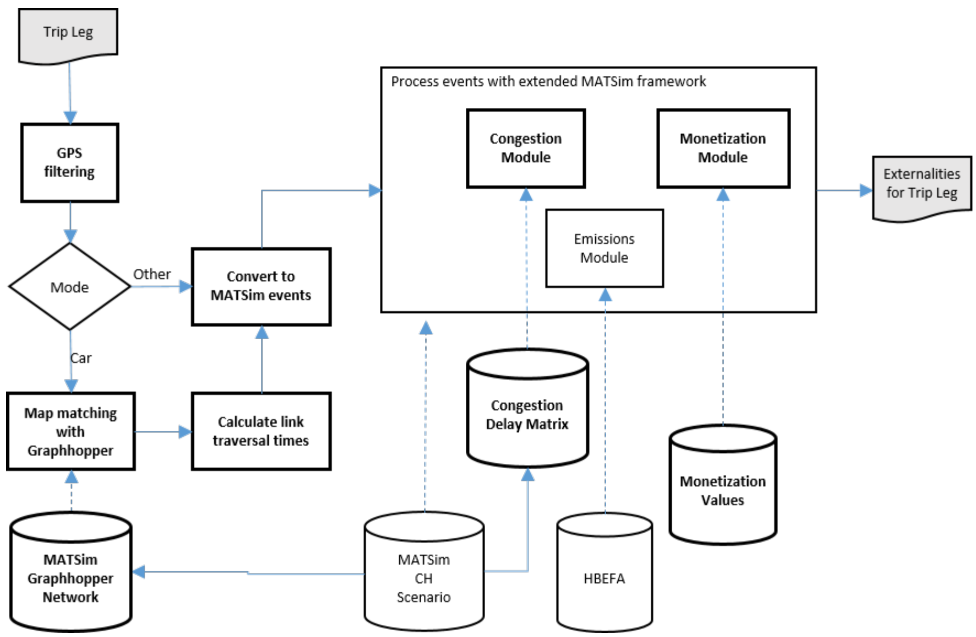
Yang, Jun, Avralt-Od Purevjav, and Shanjun Li. 2020. "The marginal cost of traffic congestion and road pricing: Evidence from a natural experiment in Beijing." *American Economic Journal: Economic Policy*, 12(1): 418–53.

Zurich Public Transport. 2017. "Servicequalität 2017- Bericht zur Servicequalität und Nachfrage auf dem Netz der Verkehrsbetriebe Zürich (german)." Zurich Public Transport (VBZ) Yearly Report.

Online Appendix

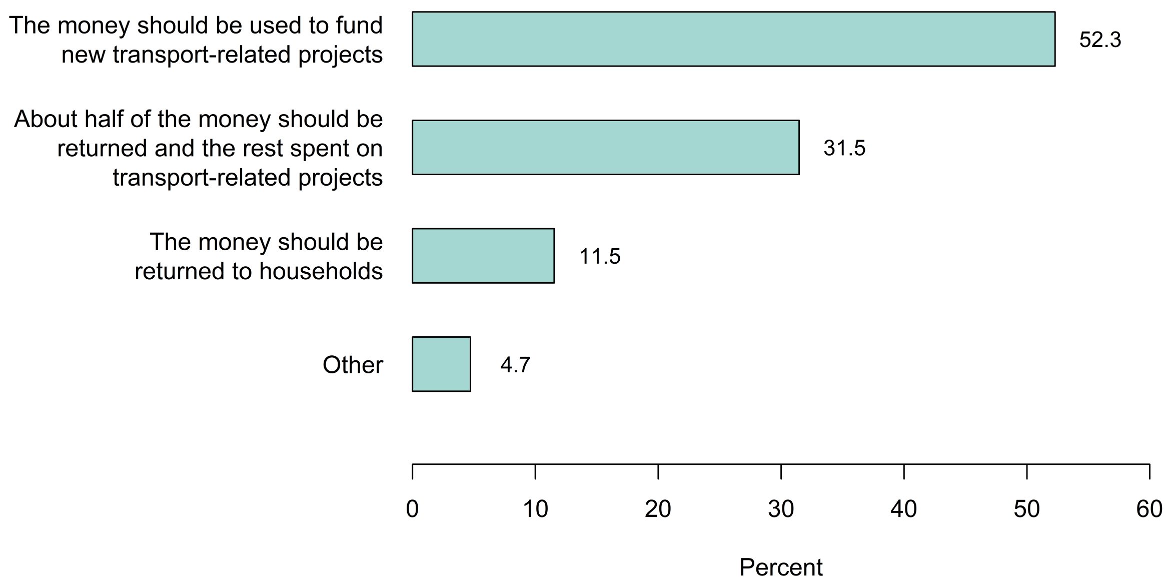
A Additional tables and figures

Figure A.1: Externalities pipeline for private motorized transport



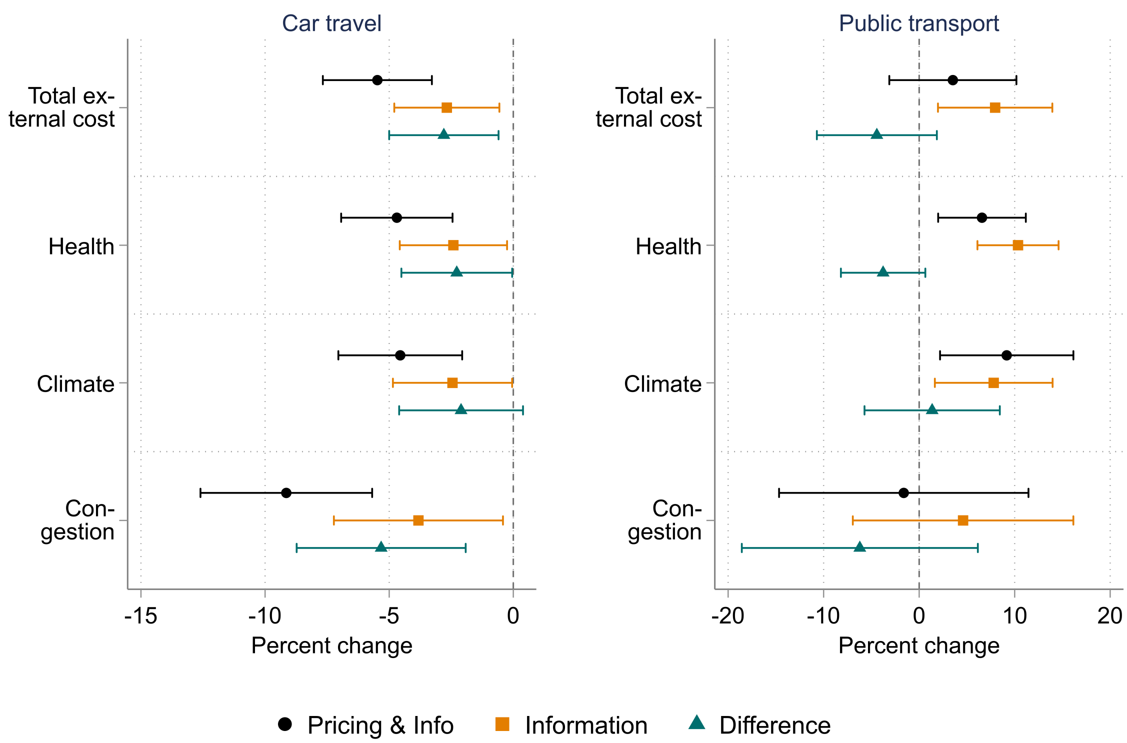
Notes: The objects in bold are those developed as part of this project. Dotted lines indicate data inputs from static sources, and solid lines are the flow of the GPS-based trip data through the model. The lack of flows inside the MATSim framework is intentional, as those modules are built on top of the MATSim event framework (Horni, Nagel and Axhausen, 2016). Molloy, Tchervenkov and Axhausen (2021) contains more details about the externalities pipeline.

Figure A.2: How revenue from transport pricing should be used.



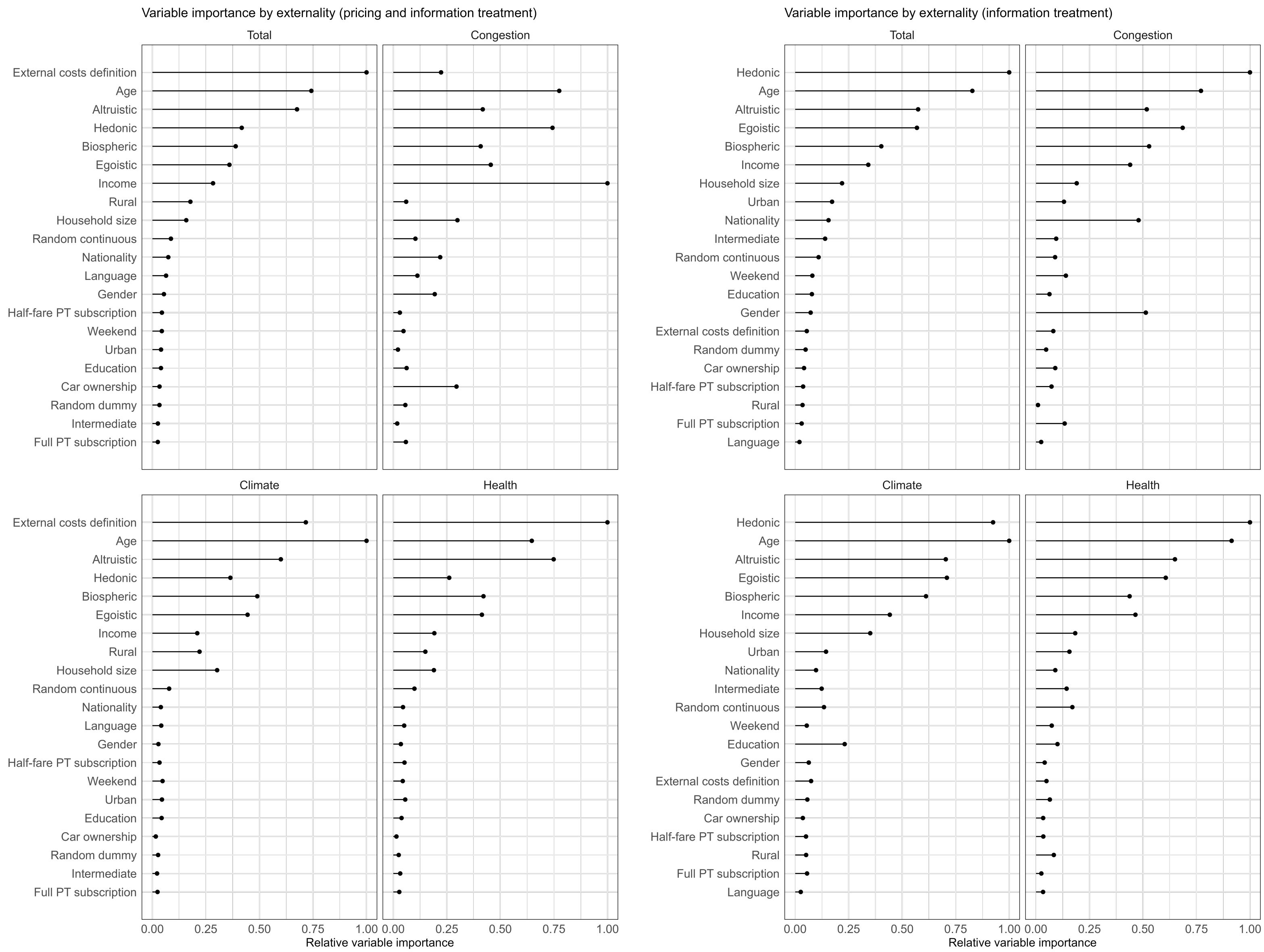
Notes: Based on question: "If dynamic mobility pricing (i.e., prices depending on mode, route and time) were introduced, what should be done with the revenue?".

Figure A.3: Treatment effect for car travel and public transport



Notes: The figure shows the proportional Average Treatment Effect for driving (left) and public transport (right). The bars show 80% confidence intervals, such that the probability mass to the right of the upper limit is 10% each (corresponding to $p < 0.1$ in a one-sided test). The results for "Difference" are the causal effect of adding pricing to existing information.

Figure A.4: Variable importance in Causal Forest



Notes: This figure shows the variable importance measure from the causal forest approach, relative to the "most important" variable, which differs across the cost dimensions considered. The variable importance measures are shown for the pricing and information treatment (left panels) as well as the information treatment (right panels).

Table A.1: Tracking summary statistics, by mode

		Car						Public transport					
		Pre-treatment			Post-treatment			Pre-treatment			Post-treatment		
Dimension	Outcome	Control	Info	Pricing	Control	Info	Pricing	Control	Info	Pricing	Control	Info	Pricing
Ext. costs	Total	4.38	4.45	4.52	4.13	4.12	4.07	0.27	0.30	0.33	0.26	0.31	0.34
(CHF)		(5.69)	(5.67)	(5.82)	(5.41)	(5.55)	(5.44)	(0.96)	(1.01)	(1.17)	(0.91)	(1.07)	(1.08)
	Congestion	0.93	0.95	0.99	0.75	0.74	0.75	0.10	0.13	0.15	0.10	0.13	0.15
		(1.44)	(1.43)	(1.50)	(1.31)	(1.34)	(1.32)	(0.71)	(0.77)	(0.89)	(0.67)	(0.82)	(0.83)
	Climate	0.86	0.87	0.88	0.84	0.82	0.82	0.02	0.02	0.02	0.01	0.02	0.02
		(1.29)	(1.29)	(1.30)	(1.24)	(1.28)	(1.23)	(0.06)	(0.06)	(0.07)	(0.05)	(0.06)	(0.06)
	Health	2.58	2.64	2.65	2.54	2.55	2.50	0.15	0.15	0.17	0.14	0.16	0.17
		(3.55)	(3.56)	(3.63)	(3.49)	(3.58)	(3.49)	(0.40)	(0.38)	(0.43)	(0.37)	(0.40)	(0.41)
Private cost		24.40	25.05	25.06	24.09	24.24	23.73	1.66	1.64	1.81	1.62	1.72	1.83
(CHF)		(33.78)	(34.15)	(34.71)	(33.56)	(34.25)	(33.43)	(4.45)	(4.22)	(4.77)	(4.37)	(4.41)	(4.58)
Tracking	Distance	34.90	35.84	35.89	34.48	34.77	34.13	9.63	9.75	11.10	8.92	10.27	10.98
	(km)	(48.25)	(48.79)	(49.71)	(47.97)	(49.00)	(47.92)	(32.44)	(30.77)	(34.83)	(30.54)	(33.08)	(33.28)
	Duration	50.20	50.98	50.89	49.31	48.81	48.63	15.13	15.30	16.36	14.79	16.86	17.06
	(min)	(54.62)	(55.82)	(55.99)	(54.30)	(55.29)	(57.33)	(47.59)	(43.63)	(45.53)	(43.95)	(48.21)	(44.89)

		Bicycle						Walking					
		Pre-treatment			Post-treatment			Pre-treatment			Post-treatment		
Dimension	Outcome	Control	Info	Pricing	Control	Info	Pricing	Control	Info	Pricing	Control	Info	Pricing
Ext. costs	Total	0.05	0.05	0.05	0.04	0.04	0.04	−0.19	−0.20	−0.20	−0.18	−0.19	−0.19
(CHF)		(0.27)	(0.25)	(0.30)	(0.21)	(0.21)	(0.25)	(0.26)	(0.26)	(0.26)	(0.24)	(0.25)	(0.25)
	Congestion												
	Climate												
	Health	0.05	0.05	0.05	0.04	0.04	0.04	−0.19	−0.20	−0.20	−0.18	−0.19	−0.19
		(0.27)	(0.25)	(0.30)	(0.21)	(0.21)	(0.25)	(0.26)	(0.26)	(0.26)	(0.24)	(0.25)	(0.25)
Private cost													
(CHF)													
Tracking	Distance	0.72	0.66	0.71	0.51	0.50	0.56	1.71	1.77	1.77	1.64	1.71	1.75
	(km)	(3.93)	(3.55)	(4.31)	(3.02)	(3.01)	(3.51)	(2.30)	(2.33)	(2.31)	(2.16)	(2.24)	(2.26)
	Duration	2.46	2.28	2.34	1.83	1.74	1.88	24.99	25.08	24.84	22.63	23.35	23.81
	(min)	(13.31)	(11.35)	(13.42)	(11.79)	(9.89)	(10.71)	(46.95)	(42.07)	(43.91)	(40.22)	(42.03)	(44.11)

Notes: Average values per participant and day during the experiment. Standard deviations in parentheses.

Table A.2: Treatment effects for car and public transport

(a) Car

	Total External Costs	Health Costs	Climate Costs	Congestion Costs
Pricing	-0.226**	-0.119**	-0.038*	-0.069**
	(0.071)	(0.044)	(0.016)	(0.020)
Information	-0.110	-0.061	-0.020	-0.029
	(0.068)	(0.043)	(0.016)	(0.020)
Difference	-0.115	-0.058	-0.018	-0.040*
	(0.071)	(0.044)	(0.016)	(0.020)
Adj R^2	0.238	0.229	0.224	0.273
Clusters	3,539	3,539	3,539	3,539
N	164,912	164,912	164,912	164,912

(b) Public Transport

	Total External Costs	Health Costs	Climate Costs	Congestion Costs
Pricing	0.009	0.009'	0.001'	-0.002
	(0.013)	(0.005)	(0.001)	(0.010)
Information	0.020'	0.015**	0.001	0.005
	(0.012)	(0.005)	(0.001)	(0.009)
Difference	-0.011	-0.005	0.000	-0.006
	(0.013)	(0.005)	(0.001)	(0.010)
Adj R^2	0.285	0.243	0.172	0.270
Clusters	3,539	3,539	3,539	3,539
N	164,912	164,912	164,912	164,912

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors in parentheses and clustered at participant level. The dependent variable contains the external costs of transport aggregated to the person-day level (in CHF). Difference is the differential effect between the Pricing and the Information groups. All regressions include individual, calendar day and day of study FE.

Table A.3: Multivariate interactions

	Total Costs			Health Costs			Climate Costs			Congestion Costs		
	Pricing	Info.	Diff.	Pricing	Info.	Diff.	Pricing	Info.	Diff.	Pricing	Info.	Diff.
Base	0.696**	0.104	0.592’	0.336*	-0.012	0.348’	0.130*	0.038	0.092	0.230**	0.078	0.152
	(0.245)	(0.239)	(0.335)	(0.152)	(0.152)	(0.210)	(0.057)	(0.056)	(0.078)	(0.072)	(0.073)	(0.100)
Male	-0.168	-0.156	-0.012	-0.057	-0.021	-0.036	-0.019	-0.020	0.002	-0.093**	-0.115**	0.022
	(0.108)	(0.099)	(0.146)	(0.068)	(0.062)	(0.092)	(0.025)	(0.023)	(0.034)	(0.033)	(0.030)	(0.045)
Income>12k	-0.113	0.003	-0.116	-0.009	0.015	-0.025	-0.010	-0.020	0.010	-0.093*	0.007	-0.101
	(0.136)	(0.132)	(0.190)	(0.082)	(0.082)	(0.116)	(0.033)	(0.030)	(0.045)	(0.046)	(0.042)	(0.062)
Income<8k	0.179	-0.008	0.187	0.064	-0.049	0.113	0.038	-0.002	0.040	0.077*	0.042	0.035
	(0.120)	(0.115)	(0.166)	(0.075)	(0.071)	(0.103)	(0.027)	(0.027)	(0.038)	(0.036)	(0.035)	(0.050)
Age>54	-0.176	0.025	-0.201	-0.108	-0.015	-0.092	-0.061’	0.003	-0.064	-0.008	0.036	-0.044
	(0.147)	(0.147)	(0.208)	(0.092)	(0.090)	(0.129)	(0.034)	(0.034)	(0.048)	(0.046)	(0.041)	(0.062)
Age<30	-0.531**	0.133	-0.663**	-0.278**	0.078	-0.356**	-0.118**	0.026	-0.144**	-0.134**	0.029	-0.163**
	(0.144)	(0.116)	(0.184)	(0.089)	(0.074)	(0.115)	(0.034)	(0.027)	(0.043)	(0.045)	(0.037)	(0.058)
Tertiary ed.	-0.074	-0.020	-0.053	-0.075	-0.027	-0.049	-0.017	-0.017	0.000	0.018	0.023	-0.005
	(0.117)	(0.112)	(0.162)	(0.071)	(0.071)	(0.100)	(0.026)	(0.026)	(0.037)	(0.039)	(0.031)	(0.050)
HH size>4	-0.234	0.200	-0.434’	-0.163	0.116	-0.279’	-0.091*	0.037	-0.129’	0.020	0.047	-0.026
	(0.181)	(0.189)	(0.262)	(0.113)	(0.107)	(0.156)	(0.043)	(0.050)	(0.066)	(0.047)	(0.056)	(0.073)
HH size<3	-0.125	-0.012	-0.113	-0.010	0.010	-0.020	-0.029	-0.024	-0.005	-0.086*	0.002	-0.088’
	(0.120)	(0.113)	(0.165)	(0.074)	(0.071)	(0.103)	(0.028)	(0.027)	(0.038)	(0.038)	(0.033)	(0.050)
French sp.	0.250*	0.026	0.224	0.125	-0.007	0.132	0.047’	-0.010	0.058	0.078*	0.043	0.034
	(0.123)	(0.116)	(0.169)	(0.079)	(0.074)	(0.108)	(0.028)	(0.027)	(0.039)	(0.036)	(0.034)	(0.049)
English sp.	0.301	0.073	0.228	0.117	0.061	0.056	0.060	0.006	0.055	0.124’	0.006	0.117
	(0.210)	(0.224)	(0.307)	(0.126)	(0.144)	(0.192)	(0.052)	(0.053)	(0.074)	(0.066)	(0.067)	(0.095)
Foreign	-0.239’	0.151	-0.390*	-0.103	0.072	-0.175	-0.040	0.040	-0.080’	-0.096*	0.039	-0.135*
	(0.140)	(0.131)	(0.192)	(0.084)	(0.083)	(0.118)	(0.033)	(0.030)	(0.044)	(0.045)	(0.040)	(0.061)
Non-urban	-0.017	0.253*	-0.270’	-0.023	0.139*	-0.162	-0.010	0.043’	-0.053	0.016	0.071*	-0.055
	(0.119)	(0.110)	(0.162)	(0.075)	(0.070)	(0.102)	(0.027)	(0.026)	(0.038)	(0.034)	(0.033)	(0.047)
Rural	-0.457*	0.085	-0.541*	-0.282*	0.037	-0.319’	-0.123*	0.008	-0.130*	-0.052	0.040	-0.092
	(0.203)	(0.183)	(0.273)	(0.130)	(0.127)	(0.182)	(0.048)	(0.043)	(0.064)	(0.058)	(0.040)	(0.071)
Car owner	-0.370*	-0.116	-0.254	-0.143	-0.008	-0.135	-0.054	-0.015	-0.039	-0.173**	-0.093’	-0.080
	(0.171)	(0.159)	(0.234)	(0.105)	(0.097)	(0.143)	(0.039)	(0.037)	(0.054)	(0.051)	(0.051)	(0.072)
1/2 fare	-0.086	-0.035	-0.051	-0.069	-0.008	-0.061	-0.021	0.004	-0.025	0.004	-0.031	0.035
	(0.106)	(0.101)	(0.146)	(0.066)	(0.064)	(0.092)	(0.024)	(0.023)	(0.034)	(0.032)	(0.031)	(0.044)
Weekend	0.150	0.209’	-0.059	0.077	0.111	-0.033	0.046’	0.027	0.019	0.027	0.071*	-0.044
	(0.121)	(0.118)	(0.147)	(0.077)	(0.074)	(0.093)	(0.026)	(0.026)	(0.032)	(0.034)	(0.036)	(0.043)
Correct EC	-0.379**	-0.083	-0.295*	-0.239**	-0.064	-0.175’	-0.082**	-0.033	-0.048	-0.058’	0.014	-0.072
	(0.110)	(0.101)	(0.149)	(0.069)	(0.063)	(0.093)	(0.026)	(0.023)	(0.034)	(0.034)	(0.031)	(0.046)
Egoistic	0.049	-0.093	0.142	0.057	-0.073	0.129	0.028	-0.042’	0.070*	-0.036	0.021	-0.057
	(0.107)	(0.109)	(0.153)	(0.066)	(0.068)	(0.095)	(0.025)	(0.025)	(0.036)	(0.033)	(0.032)	(0.046)
Altruistic	-0.151	-0.233*	0.082	-0.090	-0.125’	0.035	-0.031	-0.046’	0.015	-0.030	-0.062*	0.033
	(0.117)	(0.107)	(0.158)	(0.073)	(0.069)	(0.100)	(0.027)	(0.025)	(0.036)	(0.036)	(0.031)	(0.048)
Hedonic	0.061	-0.083	0.144	0.024	0.003	0.021	0.008	0.001	0.006	0.029	-0.087**	0.117*
	(0.114)	(0.108)	(0.157)	(0.070)	(0.068)	(0.097)	(0.027)	(0.025)	(0.037)	(0.036)	(0.032)	(0.048)
Biospheric	-0.163	0.126	-0.289’	-0.089	0.074	-0.164	-0.037	0.032	-0.069’	-0.037	0.019	-0.056
	(0.121)	(0.114)	(0.167)	(0.077)	(0.073)	(0.106)	(0.028)	(0.026)	(0.038)	(0.037)	(0.033)	(0.049)
Adj. R^2		0.232			0.224			0.222			0.268	
Clusters		3,419			3,419			3,419			3,419	
N		159,965			159,965			159,965			159,965	

Notes: **: $p < 0.01$, *: $p < 0.05$, ’: $p < 0.1$. Standard errors in parentheses and clustered at the participant level. The dependent variable is the external cost of transport aggregated to the person-day level. The “Pricing”, “Info.” and “Diff.” columns indicate the type of DiD term with which the interaction terms have been multiplied. All dimensions also include one omitted category. The “Base” coefficient is thus associated with an observation that has a zero for all included dummies. Income refers to monthly household income, in CHF. “French sp.” and “English sp.” denotes respondents who chose to answer the surveys in French and English, respectively. ”Non-urban” denotes municipalities that are not labeled as urban nor as rural by the Swiss Federal Office of Statistics.

Table A.4: ATE on travel distance

(a) Overall margin

	Distance Total	Distance Car	Distance Public Transport	Distance Bicycle	Distance Walk
Pricing	0.982	0.956*	1.050	1.112	1.044*
	(0.015)	(0.017)	(0.042)	(0.084)	(0.019)
Information	1.006	0.973	1.119**	1.053	1.012
	(0.015)	(0.016)	(0.045)	(0.074)	(0.018)
Difference	0.976'	0.983	0.938'	1.056	1.032'
	(0.014)	(0.017)	(0.034)	(0.082)	(0.018)
Precipitation	0.998*	1.000	0.992**	0.978**	0.993**
	(0.001)	(0.001)	(0.002)	(0.006)	(0.001)
Heat	1.048**	1.062**	1.004	1.041*	1.019**
	(0.004)	(0.005)	(0.010)	(0.016)	(0.004)
Cold	0.857**	0.827**	0.909'	0.893'	0.952**
	(0.017)	(0.021)	(0.046)	(0.053)	(0.017)
Adj. R^2	0.271	0.297	0.414	0.471	0.266
Clusters	3,539	3,539	3,421	2,166	3,539
N	164,912	164,912	159,545	101,584	164,912

(b) Intensive margin

	Total	Car	Public Transport	Bicycle	Walking
Pricing	0.984	0.974'	1.020	0.971	1.046**
	(0.014)	(0.015)	(0.037)	(0.050)	(0.017)
Information	1.006	0.986	1.087*	1.006	1.015
	(0.014)	(0.015)	(0.040)	(0.047)	(0.017)
Difference	0.978	0.988	0.939'	0.965	1.031'
	(0.014)	(0.016)	(0.032)	(0.053)	(0.017)
Precipitation	0.999	1.000	0.997	1.001	0.995**
	(0.001)	(0.001)	(0.002)	(0.004)	(0.001)
Heat	1.049**	1.066**	1.025**	1.023**	1.022**
	(0.004)	(0.004)	(0.009)	(0.009)	(0.004)
Cold	0.854**	0.837**	0.938	0.942	0.952**
	(0.017)	(0.019)	(0.041)	(0.041)	(0.016)
Adj. R^2	0.277	0.295	0.484	0.642	0.260
Clusters	3,539	3,539	3,271	1,555	3,538
N	158,116	128,790	55,594	13,152	136,171

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. The dependent variable contains the distance traveled aggregated to the person-day level either including zeroes (panel a) or restricted to positive observations (panel b). Standard errors (in parentheses) are clustered at the participant level. The coefficients were estimated using a PPML model and then exponentiated to derive proportional effects (with 1.00 representing no effect). All regressions include individual, calendar day and day of study FE.

Table A.5: ATE on probability to travel (extensive margin)

	Total Travel	Car	Public Transport	Bicycle	Walking
Information	0.998	0.982*	1.044’	1.005	1.002
	(0.003)	(0.009)	(0.025)	(0.058)	(0.008)
Pricing	0.999	0.971**	1.061*	1.137*	1.006
	(0.003)	(0.010)	(0.028)	(0.069)	(0.008)
Heat	0.999*				
	(0.001)				
Cold	1.002				
	(0.004)				
Precipitation	1.000**				
	(0.000)				
Difference	1.001	0.989	1.016	1.131’	1.004
	(0.003)	(0.010)	(0.026)	(0.071)	(0.008)
N	164,912	155,126	159,395	101,584	157,317

Notes: **: $p < 0.01$, *: $p < 0.05$, ’: $p < 0.1$. Standard errors in parentheses and clustered at participant level. The coefficients in the table are marginal effects in the form of semi-elasticities (dlny/dx) after estimating a logit regression that includes the weather variables and a series of dummies to capture person, day of study and calendar day FE effects.

Table A.6: ATE on departure time for car trips

	Overall	Morning	Evening
Pricing	0.510	-4.638*	1.702
	(2.543)	(1.977)	(2.185)
Information	-3.228	-1.653	0.065
	(2.520)	(2.003)	(2.216)
Difference	3.738	-2.985	1.637
	(2.553)	(2.108)	(2.251)
Adj R^2	0.052	0.204	0.116
Clusters	2,882	2,881	2,882
N	277,945	100,446	177,497

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors (in parentheses) clustered at participant level. The regressions include observations from participants that travelled at least once by car in the morning peak (departure between 6:30 and 8:30) and the evening peak (departure between 16:30 and 18:30) during the observation period. In column 1, all trips were combined, whereas columns 2 and 3 focus on departure before or after noon, respectively. All regressions include day of calendar, day of study and person fixed effects.

Table A.7: ATE on mode distance share and congestion per km

(a) Mode Distance Share

	Car	Public transport	Bicycle	Walking
Pricing	0.971**	1.042'	1.174*	1.054*
	(0.008)	(0.023)	(0.079)	(0.028)
Information	0.986'	1.040'	1.064	1.008
	(0.007)	(0.022)	(0.067)	(0.026)
Difference	0.984*	1.001	1.103	1.046'
	(0.008)	(0.021)	(0.075)	(0.027)
Precipitation	1.002**	0.995**	0.980**	0.996**
	(0.000)	(0.001)	(0.004)	(0.002)
Heat	1.005**	0.967**	0.987	1.016**
	(0.002)	(0.005)	(0.012)	(0.006)
Cold	0.970**	1.027	0.979	1.021
	(0.010)	(0.023)	(0.041)	(0.022)
Adj. R^2	0.054	0.203	0.304	0.124
Clusters	3,539	3,421	2,166	3,539
N	158,116	153,133	97,996	158,116

(b) Congestion and crowding per km

	Congestion (car)	Congestion (PT)
Pricing	0.927**	0.967
	(0.018)	(0.040)
Information	0.956*	0.934
	(0.020)	(0.039)
Difference	0.970	1.035
	(0.020)	(0.041)
Precipitation	0.999	0.997
	(0.001)	(0.002)
Heat	0.945**	0.944**
	(0.004)	(0.009)
Cold	1.080**	1.083
	(0.029)	(0.058)
Adj. R^2	0.028	0.038
Clusters	3,539	2,433
N	128,790	47,956

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors in parentheses and clustered at participant level. In panel (a), the dependent variable is the share of each mode per person and day (between 0 and 1); in panel (b), the dependent variable are the external congestion costs per km of either car or PT travel. Both models are estimated by PPML and include Person, day of calendar and day of study FE.

Table A.8: Trends in distance travelled

	All Travel	All Travel	Car	PT	Bicycle	Walking
Info	1.120	0.697	0.176	0.461	-0.052	0.113’
	(1.109)	(1.452)	(1.322)	(0.834)	(0.117)	(0.061)
Pricing	2.484*	2.611’	0.910	1.561’	0.032	0.108’
	(1.134)	(1.496)	(1.333)	(0.904)	(0.136)	(0.059)
Study Day	0.038	0.031	-0.040	0.065*	-0.001	0.007**
	(0.038)	(0.052)	(0.045)	(0.031)	(0.004)	(0.002)
Pricing x Study Day		-0.008	0.003	-0.005	-0.003	-0.004
		(0.062)	(0.055)	(0.036)	(0.005)	(0.003)
Info x Study Day		0.028	0.052	-0.020	-0.001	-0.003
		(0.062)	(0.055)	(0.036)	(0.004)	(0.003)
Calendar day FE	☑	☑	☑	☑	☑	☑
Adj. R^2	0.005	0.005	0.005	0.007	0.007	0.014
N	83,855	83,855	83,855	83,855	83,855	83,855

Notes: **: p < 0.01, *: p < 0.05, ’: p < 0.1. The dependent variable the travel distance (in km), aggregated to the person-day level. Standard errors (in parentheses) are clustered at the participant level.

B The MOBIS study: Step by step

B.1 Sample size

In order to determine the appropriate sample size of the experiment, we carried out a series of power calculations by means of simulation. In panel data, autocorrelation is a design feature, which we also observe in our data (i.e., a particular respondent makes similar travel choices over time). The presence of autocorrelation implies that the standard formulae for power calculations, e.g. as in Duflo, Glennerster and Kremer (2007), are biased (Burlig, Preonas and Woerman, 2020). Computing the power of an experiment based on simulations addresses this problem as it uses the empirical correlation structure in the data.

We based our power calculations on data from two earlier transport studies carried out by ETH-IVT.[24] We imposed a significance level of p=0.05, a power of 0.8 and an effect size of 5%. Given these settings, the power calculations indicated that we needed a sample size of around 1,100 for each group (treatment and control). Given that we have two treatment groups, this led to a target sample size of 3,300 for our study. To ensure that this sample size was attained even after removing respondents who did not participate on a sufficient number of days or who had to be excluded for other reasons, we set a recruitment goal of 3,600 people. Once we attained this number, recruitment was stopped.

B.2 Invitation

Assuming a participation rate in the main study similar to the one observed in the pilot study (i.e. 3.4%), approximately 100,000 addresses were expected to be required to achieve the goal of 3,600 participants. Ultimately, 90,909 persons were invited to participate in the MOBIS study. The invitation letters were sent in two waves by regular mail.

The first wave started in July 2019. 60,409 persons were contacted using home addresses provided by the Swiss Federal Statistical Office (BFS in German). Only persons aged 18 to 65 in 2018 and living in an agglomeration of the German- and French-speaking Swiss cantons were invited. The persons who did not react after the first invitation letter received up to two reminders (a second and third invitation letter). The time lag between the invitation letter and the reminders was between 3 and 4 weeks.

[24] The 6-weeks MOBIDrive (Axhausen et al., 2002) and the 6 week-Thurgau survey (Axhausen et al., 2007).

The second wave was invited in October 2019. 30,500 additional persons were contacted using home addresses purchased from the private vendor Schober Information Group AG. The persons of the second wave only received a single invitation letter, i.e., no reminder was sent. 56 persons in the second wave accidentally received a duplicated participant ID which had already been allocated to participants in the first wave. These persons were informed that despite the invitation, they could not participate in the study.

The invitation letter was written in German, French and English. The front side of the invitation showed the German or the French version, while the back side always showed the English version. The language of the front side was assigned based on the communication language, which was provided in the list of addresses. In case of Italian speaking persons, the main spoken language of their home canton was assigned.
The content of the invitation was the same for all languages. The letter explained the background of the study (rationale, participating universities and supporting institutions) and provided instructions for completing the online introduction survey and registration. No reference to an experiment was made. The participant ID (a five-letter code) was provided in these instructions. This ID enabled access to the introduction survey, registration and final survey. Finally, the letter informed about the financial reward for complete participation and about the data privacy policy.

B.3 Introduction survey

The invited persons who were willing to participate in the MOBIS study first had to fill out an online introduction survey. This survey had two goals: First, to collect transport-related opinions from the general population, and second, to identify subjects who qualified for the main study based on the following inclusion criteria:

- Be the recipient of the personal invitation letter (the invitation was not transferable to other persons)
- Live in a metropolitan area in the German- or French-speaking part of Switzerland (the lists of addresses included only people living in these areas but the survey double-checked the post code)
- Be between 18 and 65 years old in 2018 (the list of addresses provided by the BFS was pre-filtered by age at this year)
- Travel by car at least two weekdays per week (including their own car, car-sharing as a driver, or with a taxi and App-based services such as Uber as passenger)

- Use a smartphone that can install the tracking app
- To be able to walk 200 meter without assistance (to ensure that participants have free mode choice)
- Not work as a professional driver (to ensure that participants have free mode choice)

People meeting the above listed inclusion criteria were invited to register for the field experiment of the MOBIS study by clicking on a web link embedded at the end of the introduction survey survey. Beyond the questions related to the aforementioned inclusion criteria, the introduction survey contained questions related to transport-related topics, but without making reference to an experiment.

The introduction survey was accessible online through the survey platform Qualtrics.

B.4 Tracking

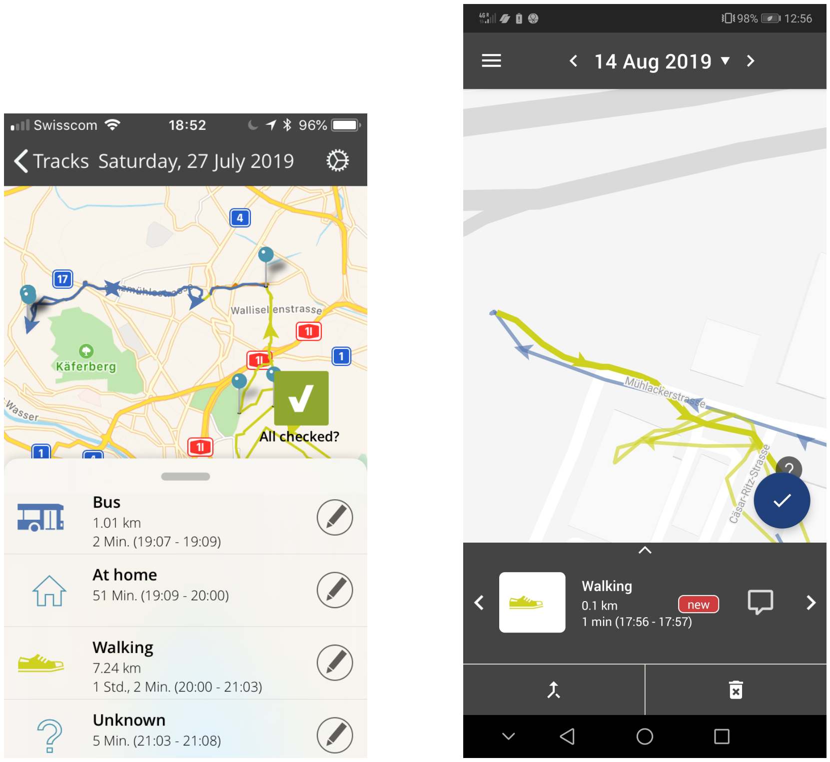
People who completed the introduction survey, met the inclusion criteria and accepted to participate in the tracking study were sent forward to the registration (survey), which was also accessible online through Qualtrics. The registration had as a goal to confirm the acceptance of the conditions and terms of the field experiment and to collect the emails of the participants for sending interventions (incentives) and information. All registered participants were given a code and directions to install the Catch-my-Day app.

Participants could start tracking at any time, and the 8 weeks would start from the first complete tracking day. To remain eligible for the CHF 100 incentive, participants were informed that they needed to track at least half the time for the duration of the study. Participants identified as not tracking for a certain number of consecutive days were notified by email, with the aim of increasing the quality of the tracking data and reducing the dropout rate. An initial minimum number of 2 days between reminders was set but later increased to 4 days, to avoid burdening the participants. Participants who did not generate tracking data on more than 12 days in the first 4 weeks were not allowed to participate in the treatment phase and did not receive the CHF 100 incentive.

Users could view their daily travel patterns on their phone in the form of a logbook, validate the travel mode and activity purpose or indicate if a stage or activity did not take place. There are some user-interface differences between the iOS and Android versions, which are most noticeable in the validation interface. Figure B.1 presents the validation interface of the app for the respective operating systems.

To assess the mode detection performance of the app, we assume that a mode has been assigned correctly if (i) a stage has been validated but not corrected and (ii) the person

Figure B.1: Validation interface for iOS (left) and Android (right)



in question has made at least one correction during the course of the experiment.[25]

Table B.1 provides the accuracy rates using this assumption. There is a small difference in accuracy between iOS and Android, with iOS being on average slightly better (92.23% vs 92.10%). The differences in accuracy are more observable at the categorical level. The iOS performs better on car, local rail, regional rail, tram and walk. However, the differences are only 1-3% in accuracy. Note that 'Rail' groups all rail modes together for conciseness. It is also worth noting that while the accuracy of some individual rail modes is quite low, the overall rail accuracy is very good.

Table B.2 presents the confusion matrix between the modes. Here we can see that the algorithm often mis-detected car travel as bus travel. The category "Other *" includes those modes which could be manually selected by the participant, but which were not automatically detected. These included Carsharing, Taxi/Uber, Motorbike/Mopeds, and

[25] Among the participants that used the validation functionality (85.7%), 20.4% of iPhone users and 44.1% of Android users did not make a single correction over the 8 weeks of the experiment. Even with state-of-the-art accuracy rates, a perfect mode detection is highly unlikely. We therefore assume that these participants did not use or understand the validation interface correctly and and these participants were therefore removed from the following analysis on the mode detection performance. The difference between the two operating systems also indicates that the iPhone validation interface was much more intuitive.

Table B.1: Comparison of the mode detection performance bewteen iOS and Android

Mode	% Correct Android		iOS	
Airplane	99.48%		98.86%	
Bicycle	81.59%		79.14%	
Bus	66.98%		66.82%	
Car	92.98%		93.15%	
Rail	89.50%		91.05%	
Local train		88.67%		90.18%
Regional train		71.35%		73.40%
Subway		93.56%		92.53%
Train		63.13%		63.78%
Tram	95.01%		96.64%	
Walk	95.56%		97.21%	

Table B.2: Confusion matrix of mode detection accuracy

Predicted	Confirmed mode: Airplane	Bicycle	Boat	Bus	Car	Rail	Tram	Walk	Other	**Total**
Airplane	**2,113**	-	-	-	22	-	-	-	-	**2,135**
Bicycle	4	**26,201**	136	438	1,499	177	149	2,771	1,500	**32,875**
Bus	1	435	2	**35,713**	15,085	140	280	889	865	**53,410**
Car	372	2,495	741	8,028	**366,649**	3,314	1,950	2,834	7,433	**393,816**
Rail	64	56	85	1,748	7,298	**60,270**	691	258	298	**70,768**
Tram	-	49	2	128	396	60	**20,174**	149	16	**20,974**
Walk	80	3,807	456	1,224	9,960	868	868	**514,944**	638	**532,845**
	2,634	**33,043**	**1,422**	**47,279**	**400,909**	**64,829**	**24,112**	**521,845**	**10,750**	**1,106,823**

Gondolas. Most of these were detected as car travel, and the 1,500 'Bicycle' trips which were corrected to 'Other' were predominately trips by motorbike or moped.

B.5 Final survey

Upon completing the tracking study, participants received an email with a link to the final survey. The final survey contained questions related to the following topics:

- Socio-economic background: self-reported absences during the field experiment.
- Employment background: same questions as in introduction survey to check changes during the experiment and flexibility in working conditions regarding home office and work schedule.

- Transport-related opinions
- Awareness and evaluation of the interventions
- Opinion regarding the use of the revenues of mobility pricing
- Lifestyles and values
- Bank data for the payment

B.6 Compensation

All participants who completed the final survey received CHF 100 for their full participation, except those who did not generate tracking data on more than 12 days during the treatment phase, who instead received CHF 50 for partial participation (this partial compensation was not discussed ex-ante). Participants who did not generate enough tracking data in the observation phase were removed from the study, and thus did not receive any compensation. In addition, participants in the pricing group received any positive amount remaining on their virtual mobility budget.

Importantly, all participants were informed about the incentive of CHF 100 upon completion of the study. The possibility of a partial incentive was not mentioned and introduced at the end mainly as a gesture of appreciation towards people that delivered some tracks (but not enough to be included in the study). Likewise, the possibility of earning money during the pricing treatment was only communicated to the pricing group, and only on day 29 of participation.

A form was provided at the end of the final survey in which the participants could enter their bank account details, and all payments were processed by the ETH finance department. Table B.3 shows a summary of the allocated virtual budgets, remaining balances paid out to the participants as well as the incurred costs. Only the 1,147 participants who completed the pricing treatment and received compensation are included. Remaining balances (i.e., exhausted budget) are capped to zero, as this is the amount that was actually paid out. This was the case for 202 participants.

B.7 Study monitoring and user support

Two dashboards were developed for the monitoring of both the participants and the participation rate (see Figures B.2 and B.3 respectively). The first dashboard was essential for troubleshooting with participants, as it gave a visual overview of their participation by

Table B.3: Virtual budgets, remaining balances and incurred costs (CHF).

	Virtual budget	Remaining balance	Incurred costs
Mean	173.82	45.45	132.89
Std. dev.	101.63	48.53	81.66
Min	50.00	0.00	0.00
25%	100.00	7.00	75.72
50%	150.00	31.44	115.37
75%	230.00	68.53	172.72
Max	745.00	432.68	616.08

week, including when they track abroad. The second gave an overall view of the response rate. This helped identify that a second invitation wave was required to meet the target number of participants.

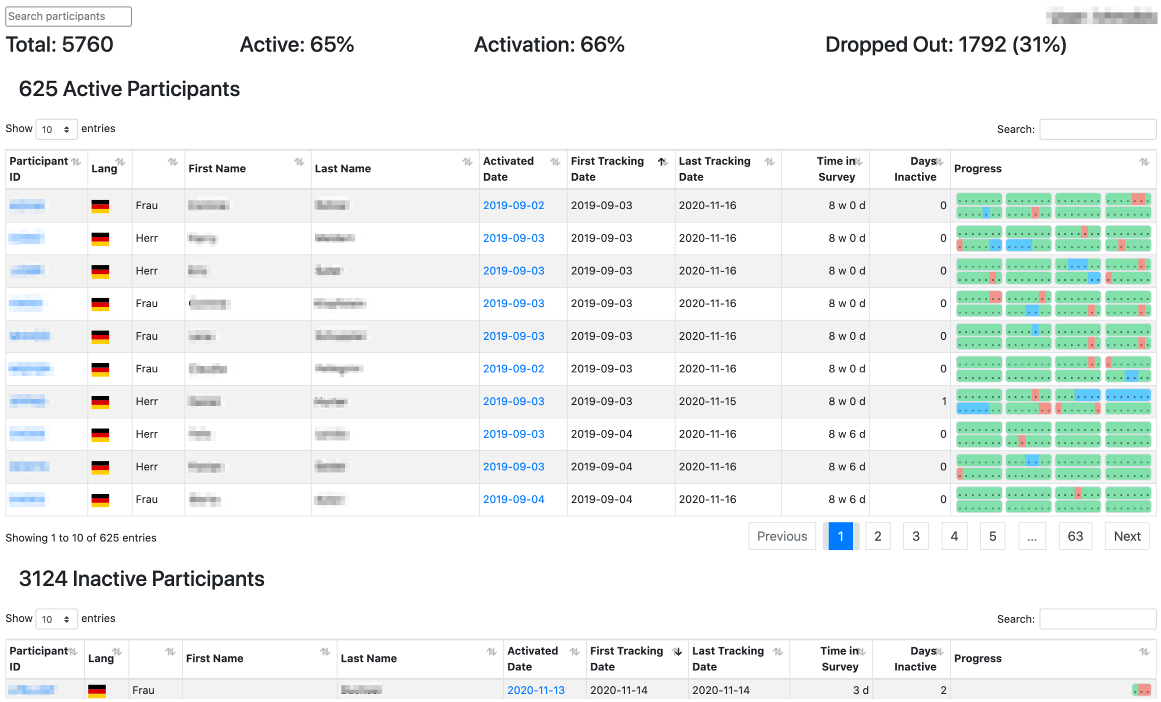
A project website was created to support people invited to the MOBIS study. The website contained links to the introduction survey and the tracking study registration, a project description, information for study participants (including a general information sheet, instructions for the tracking app, data privacy policy and consent) as well an FAQ section. The website was available in English, German and French.

Additionally, a help-desk service was set up to allow participants to ask questions and communicate any issues they might have had during the study. The communication with the help-desk was possible via phone call or email. The phone help-desk was open 10 hours per week, from 17:00 to 19:00 from Monday to Friday and from 10:00 to 12:00 on Saturday. The online help-desk received 5,218 emails during the study, of which nearly 50% came during the on-boarding process.

B.8 Participant retention

To explore the retention rate of participants in the tracking phase, we performed a survival analysis on the duration of tracking in the study. First, a Kaplan-Meier approach (see Figure B.4) shows the impact of the treatment on the length of time which participants would track. Participants who were automatically dropped out after phase 1 due to poor tracking compliance but were still tracking at the end of phase 1 were censored (marked by a cross). There is no significant difference between the three treatment groups in their survival curves. A sharp decrease in survival is evident in the last study week. As participants were informed at the end of the study that they could delete the app, the last few days of tracking were sometimes not collected before the app was deleted.

Figure B.2: Overview page of participants



Notes: This screenshot was taken after the conclusion of the study, and the participants counts do not reflect the real status during the study.

Figure B.3: Screenshot of the MOBIS response rates dashboard

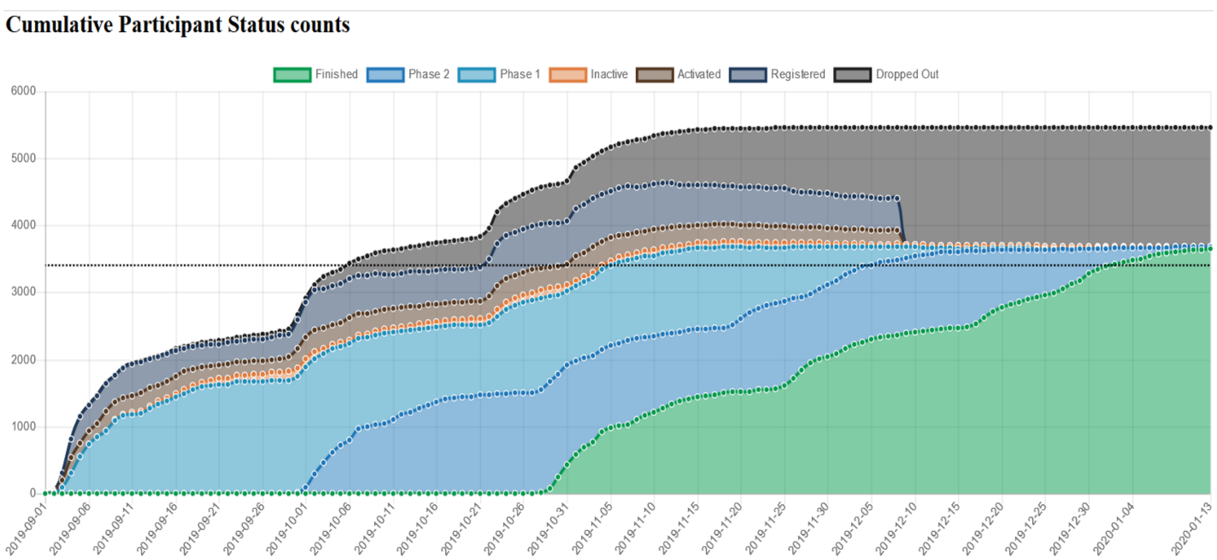
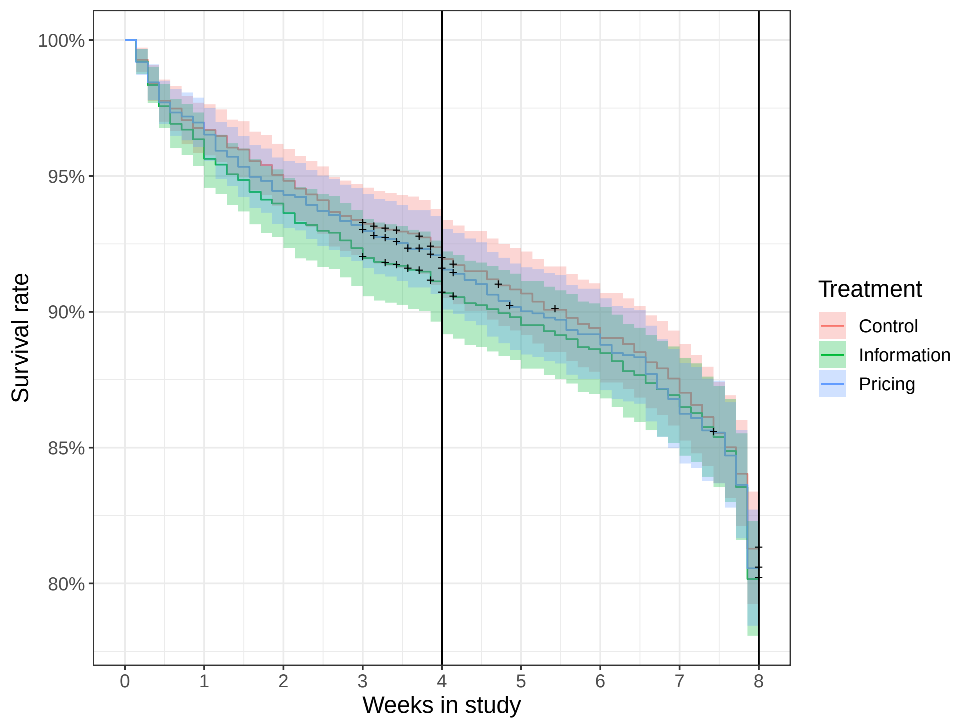


Figure B.4: Kaplan-Meier survival curve by treatment group



Notes: The cross indicates censoring of participants.

Given the participation goal of 8 weeks, one would expect that the attrition rate would be highest early on in the study and flatten out as participants neared the 8-week goal, after which they would receive the incentive. However, the survival curve is almost linear. Furthermore, Figure B.4) shows that the treatment didn't affect the attrition rate in the second phase.

A time-variant Cox proportional hazards model is to investigate the impact of different factors on the participation duration (see Table B.4 for the model results). To account for time-dependent effects, the study period was stratified into fortnightly windows. Those in high-income brackets (more than 12,000 CHF/year) were more likely to stop tracking. Conversely, those from larger households and those with tertiary education were more likely to track for longer. A significant gender-based difference was only observed in the final fortnight, where females were more likely to remain in the study.

Contrary to expectations, there was no significant effect of age on the hazard rate. This suggests that common concern about the feasibility of tracking studies for older age groups is unfounded, at least up to the age of 65, the age limit in this study.

The coefficient on employment is also time-dependent. Those in the workforce (i.e. excluding students, homeworkers and retirees) were more likely to remain in the study throughout the first fortnight.

The participant's mobile device played a much larger role. Having an Android phone of any model increased the hazard drastically. However, this effect was strongest in the first week. The effects were even larger for Huawei models. The incompatibility of GPS loggers with Android (and particularly Huawei devices) is already well known; however, here the effect is quantified, and seen to be dramatic. The effect was also time-dependent, with the most significant hazard in the first fortnight. At the end of the second fortnight, participants who tracked insufficiently were removed from the study - this explains the reduction in the Android hazard coefficient for the third fortnight, when many of them could have been expected to stop tracking, had they not been removed from the study.

At the end of the tracking study, participants were told that they could delete the app, but were also encouraged to continue using it if they wished. Figure B.5 shows the dropout rate for the whole study, including the post-study period. The majority of the participants dropped out soon after the study, but even 6 months after the study was completed, around 5% of participants continued to use the app. Anecdotal reports from participants indicated that they enjoyed having an overview of their travel, and that it even continued to inform their mobility decisions. The impacts of the mobile operating system continued even after the study, with the post-study retention rate falling faster for Android users.

Figure B.5: Post-study participation survival curve

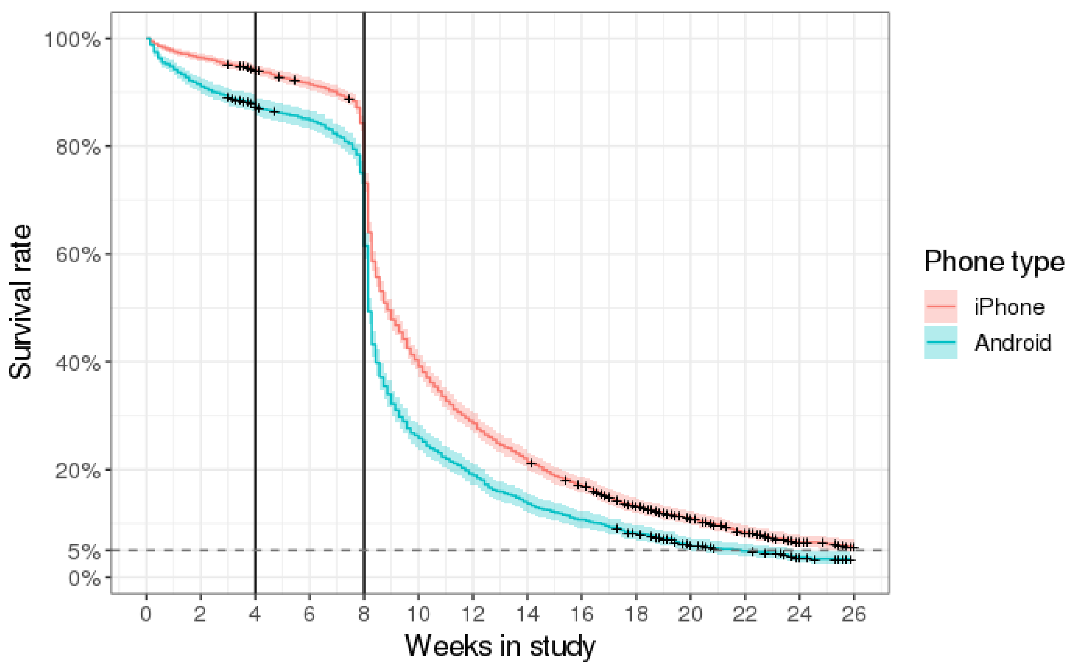


Table B.4: Cox porportional-hazard model

	Beta (SE)	HR (95% CI)	p
Income > 12,000 CHF	0.28 (0.09)	1.32 (1.10, 1.58)	0.003 **
Household size	-0.07 (0.03)	0.93 (0.87, 1.00)	0.038 *
Age (decades)	0.00 (0.03)	1.00 (0.95, 1.06)	0.883
Tertiary education	-0.19 (0.08)	0.83 (0.70, 0.97)	0.022 *
German speaking	0.03 (0.09)	1.03 (0.87, 1.22)	0.752
Female			
fortnight=1	0.02 (0.15)	1.02 (0.77, 1.35)	0.895
fortnight=2	-0.07 (0.20)	0.93 (0.62, 1.39)	0.721
fortnight=3	-0.04 (0.22)	0.96 (0.62, 1.48)	0.841
fortnight=4	-0.28 (0.12)	0.76 (0.60, 0.96)	0.022 *
Android			
fortnight=1	0.87 (0.16)	2.38 (1.73, 3.26)	0.000 ***
fortnight=2	0.46 (0.22)	1.58 (1.02, 2.45)	0.040 *
fortnight=3	-0.01 (0.25)	0.99 (0.60, 1.62)	0.960
fortnight=4	0.41 (0.13)	1.51 (1.17, 1.94)	0.002 **
Huawei			
fortnight=1	0.38 (0.20)	1.47 (0.99, 2.18)	0.057 .
fortnight=2	0.37 (0.32)	1.45 (0.78, 2.70)	0.239
fortnight=3	0.29 (0.41)	1.33 (0.59, 2.98)	0.487
fortnight=4	0.15 (0.21)	1.16 (0.77, 1.75)	0.465
Employed			
fortnight=1	-0.33 (0.16)	0.72 (0.53, 0.97)	0.033 *
fortnight=2	-0.07 (0.23)	0.94 (0.60, 1.47)	0.775
fortnight=3	0.24 (0.27)	1.27 (0.75, 2.15)	0.369
fortnight=4	0.05 (0.14)	1.05 (0.80, 1.38)	0.718
AIC		10484.33	
Coordance		0.602	
Num. events		655	
PH test		0.76	
Note:		$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$	

References

Axhausen, Kay W., Andrea Zimmermann, Stefan Schönfelder, Guido Rindsfüser, and Thomas Haupt. 2002. "Observing the rhythms of daily life: A six-week travel diary." *Transportation*, 29(2): 95–124.

Axhausen, Kay W, Michael Löchl, Robert Schlich, T Buhl, and Paul Widmer. 2007. "Fatigue in long-duration travel diaries." *Transportation*, 34(2): 143–160.

Burlig, Fiona, Louis Preonas, and Matt Woerman. 2020. "Panel data and experimental design." *Journal of Development Economics*, 102458.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using randomization in development economics research: A toolkit." *Handbook of development economics*, 4: 3895–3962.

Horni, Andreas, Kai Nagel, and Kay W Axhausen. 2016. *The multi-agent transport simulation MATSim.* Ubiquity Press, London.

Molloy, Joseph, Christopher Tchervenkov, and Kay W Axhausen. 2021. "Estimating the external costs of travel on GPS tracks." *Transportation Research Part D: Transport and Environment*, 95: 102842.