University of Basel

Faculty of
Business and Economics

October 2021

Pigovian Transport Pricing in Practice

WWZ Working Paper 2021/11

Beat Hintermann; Beaumont Schoeman; Joseph Molloy; Thomas Götschli; Alberto Castro; Christopher Tchervenkov; Uros Tomic; Kay W. Axhausen

A publication of the Center of Business and Economics (WWZ), University of Basel.

Universität Basel
Peter Merian-Weg 6
4052 Basel, Switzerland
wwz.unibas.ch

Corresponding Author:
Prof. Dr. Beat Hintermann
Tel. +41 61 207 33 39
b.hintermann@unibas.ch

Pigovian Transport Pricing in Practice*

Beat Hintermann[†] Joseph Molloy[‡] Beaumont Schoeman[†]
Thomas Götschi[†§] Alberto Castro[†] Christopher Tchervenkov[‡]
Uros Tomic[†¶] Kay W. Axhausen[‡]

September 3, 2021

Abstract

Pigovian transport pricing was implemented in a large-scale field experiment in urban areas of Switzerland. The pricing varied across time, space and mode of transport. One third of the participants were given a financial incentive to reduce their external costs of transport, whereas others were provided information only or served as a control group. The pricing treatment caused a significant reduction in the external costs of transport. This reduction is a consequence of mode substitution and a shift of departure times. The effect of providing information in the absence of pricing was statistically significant only for subgroups of the sample.

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

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

---

*This research has been supported by the Innosuisse "SCCER CREST/Mobility Joint Activity" and the Swiss Federal Roads office (grant no. 0820001327).

[†]University of Basel, Faculty of Business and Economics, Peter-Merian-Weg 6, 4002 Basel, Switzerland. Corresponding author: *b.hintermann@unibas.ch*.

[‡]ETH Zürich, Institute for Transport Planning and Systems, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland.

[§]Univ. of Oregon, School of Planning, Public Policy and Management, 97403 Eugene, OR, USA.

[¶]ZHAW, Institute of Sustainable Development, Technoparkstr. 2, 8400 Winterthur, Switzerland.

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 can be very costly as transport infrastructure competes with other land uses. Furthermore, greater road capacity induces demand and thus results in little improvement in 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 an increase 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 thus 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 uses the design of a randomized controlled trial (RCT). The sample consists of people living in urban areas of the German- and French-speaking parts of Switzerland. The pricing affects all modes and is implemented by providing the participants with a budget, from which the external costs of their transport choices is subtracted. This treatment causes a reduction in the external costs of transport by around 5.1%. If we relate this to the (total) price increase implied by the Pigovian pricing scheme of 16.4%, we arrive at an elasticity of -0.31. The reduction in the external costs is caused by a combination of a shift in transport mode and departure times, but not a reduction in the overall distance traveled. Women and French speakers respond less strongly to the pricing. Moreover, the strength of the effect is driven by intrinsic values (as measured using a standardized questionnaire), the extent to which people paid attention to the experiment and/or had prior knowledge about the external costs of transport.

[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).

To differentiate the pricing effect from a pure information effect, the experiment includes a second treatment 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*, but this effect was smaller than that of the pricing treatment and statistically significant at conventional levels only for subgroups of the sample. The results further imply that adding pricing to information is particularly important in reducing the external costs associated with congestion.

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.

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).

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 suggests that that they were effective in reducing congestion and air pollution (Isaksen and Johansen, 2021). There are two other RCT's 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 of social acceptability, behavioral change could also be achieved by means of non-financial interventions, which could be easier to implement than prices or taxes. A number of studies have investigated the effect of non-financial interventions in the transport sector Möser and Bamberg (for a review, see 2008), and some recent papers have used tracking apps to test the effect of informational interventions but 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 thus 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

The sample for the Mobility in Switzerland (MOBIS) project was recruited from individuals living in urban areas in the German- and French-speaking parts of Switzerland. Participants had to fulfill a number of inclusion criteria such as being between 18 and 65 years old and using a car at least two days per week. They 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.[2] The participants in the pricing treatment had the opportunity to earn additional money by not spending all of their budget.

Figure 1 provides an overview of the study design. We contacted over 90,000 people by mail and invited them to participate in the study. The addresses were randomly selected and provided by the Federal Office of Statistics, which maintains a comprehensive registry of inhabitants, and also from a private vendor.[3] The first step consisted in an initial online survey, which was filled in by 21,800 respondents. It contained questions about travel behavior and socio-demographics and served as a screening mechanism. Around 11,000 of the respondents from the initial survey qualified for the tracking study, and 5,466 registered for it. 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 successfully completed the RCT. The MOBIS sample differs from the general population due to an age restriction and the focus on drivers but is otherwise quite representative.

[2] At the end of the RCT (December 2020), one Swiss Franc (CHF) corresponded to 0.91 Euro or 1.11 US Dollars.

[3] 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.

Figure 1: Design of the MOBIS experiment

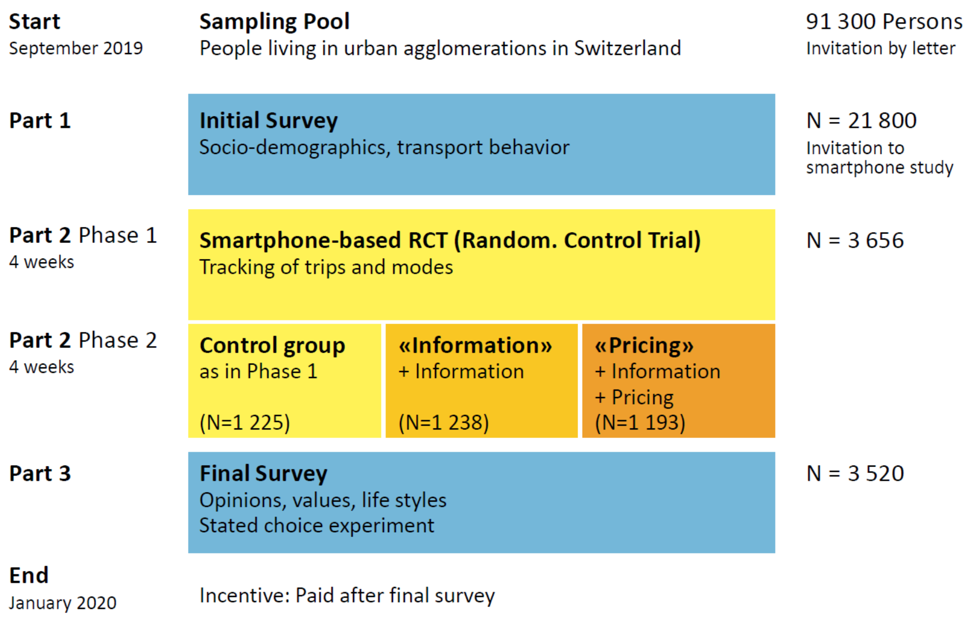


Table 1 shows the socio-demographic characteristics of our participants of the introduction survey and the tracking study and compares them to the transport Microcensus, which is a representative travel diary survey of the Swiss population undertaken by the Federal Office of Statistics and the Swiss Federal Office of Spatial Development (2017). The respondents of the MOBIS introduction survey differ from the Microcensus population in terms of the age distribution, as we limited the study to ages 18-65, and in terms of the regional coverage (only urban agglomerations and excluding the Italian-speaking region of Switzerland). The MOBIS participants have higher levels of education, employment and income, much of which was due to the age restriction.[4]

The tracking sample differs from the introduction survey sample in terms of employment, household size, income, and access to a car, probably due to the eligibility requirement of regularly traveling by car. The distribution across the language regions reflects their national representation. 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,

[4]When censoring for age, the employment rate in the Microcensus is very similar to that in MOBIS.

Table 1: Demographic information for the MOBIS sample

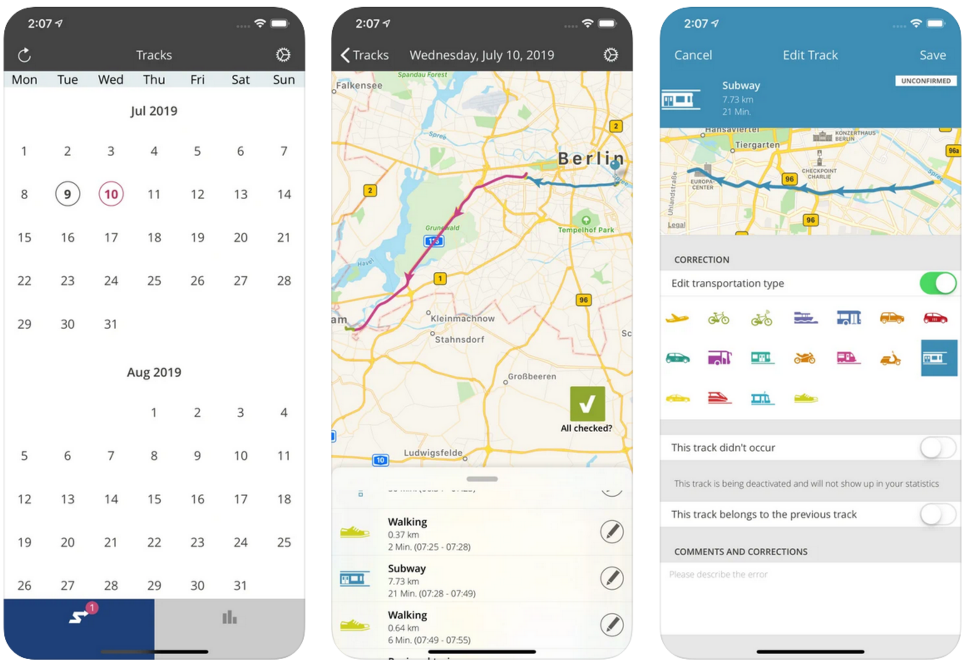
Variable	Level	MOBIS Intro	Mobis Tracking			Microcensus
			Control	Info	Pricing	
Age	Under 18					13.2
	[18, 25]	19.4	18.2	20.0	19.6	9.0
	(25, 35]	18.7	18.1	18.5	16.7	14.2
	(35, 45]	19.2	22.0	20.8	24.3	15.4
	(45, 55]	20.8	22.5	23.6	22.2	16.7
	(55, 65]	18.3	18.0	16.0	15.7	12.9
	66 and older	3.6	1.1	1.1	1.5	18.5
Education	Mandatory	9.1	8.0	5.2	6.8	19.3
	Secondary	43.6	47.6	49.4	48.4	49.5
	Higher	47.3	44.4	45.5	44.8	31.2
Employment	Employed	66.4	72.7	71.3	70.0	48.2
	Self-employed	7.4	6.1	5.4	7.4	7.2
	Apprentice	1.8	1.9	1.6	1.7	2.6
	Unemployed	4.3	3.3	4.0	4.5	2.5
	Student	9.0	7.5	8.6	7.6	3.7
	Retired	3.3	3.5	2.7	3.4	19.3
	Other	7.7	5.0	6.4	5.4	16.5
Gender	Female	50.6	50.0	50.0	50.5	50.7
	Male	49.4	50.0	50.0	49.5	49.3
Household size	1	15.8	11.3	11.6	12.2	34.0
	2	33.0	30.1	31.3	29.2	35.4
	3	20.0	23.1	21.3	20.0	13.0
	4	22.8	25.1	27.7	29.3	12.5
	5 or more	8.3	10.3	8.1	9.3	5.1
Income	4 000 CHF or less	12.1	6.8	8.2	7.2	17.8
	4 001 - 8 000 CHF	29.9	31.3	30.0	27.4	32.8
	8 001 - 12 000 CHF	24.4	27.5	29.5	30.0	17.4
	12 001 - 16 000 CHF	11.9	15.6	13.9	14.3	6.8
	More than 16 000 CHF	7.9	9.8	9.4	10.5	4.5
	Prefer not to say	13.8	9.1	9.0	10.6	20.7
Language	German	63.1	66.8	65.0	66.6	68.4
	French	28.5	25.5	26.7	26.1	25.3
	Italian					6.3
	English	8.5	7.8	8.2	7.3	
Nationality	Switzerland	98.2	98.1	98.0	98.1	75.9
	Other	1.8	1.9	2.0	1.9	24.1
Access to car	Yes	61.5	87.3	87.8	88.2	75.8
	Sometimes	15.1	11.6	10.7	10.9	18.1
	No	23.4	1.1	1.5	0.9	6.2
Full PT subscription	Yes	36.8	21.8	24.9	25.3	
	No	63.2	78.2	75.1	74.7	
Half fare PT subscription	Yes	47.9	49.2	49.0	48.3	
	No	52.1	50.8	51.0	51.7	
No PT subscription	Yes	26.0	33.9	32.6	33.9	
	No	74.0	66.1	67.4	66.1	
Access to bicycle	Yes	68.1	72.9	71.8	69.4	
	Sometimes	4.0	4.5	5.5	3.8	
	No	27.8	22.6	22.7	26.8	
N		21.571	1.225	1.238	1.193	57.090

Notes: Descriptive statistics shown for the MOBIS introduction survey (N = 21,571), the MOBIS control group (N = 1,225), the MOBIS information group (N = 1,238), the MOBIS pricing and information group (N = 1,193), and the Swiss Microcensus 2015 (N = 57,090) samples.

which uses the location services of the respective operating system. The GPS tracks are stored on the phone and uploaded to the Motiontag analytics platform, where trip stages are identified and travel modes and activities are imputed. A trip between two activities can consist of multiple stages, using the same or different modes of transport. The splitting of a day into stages and activities is called *segmentation*. This is performed by the Motiontag app using 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: The Catch-my-Day interface



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

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.[5] Around 29% of the stages were confirmed by the participants. The database stores both their correction and the original algorithmic

[5]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.

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 5.6.

The following modes are detected the by Catch-my-Day app: Air plane, bicycle, bus, 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 not detectable or 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. As seen in the tables containing the average detection accuracy by mode (Table B.1) and the mode detection matrix (Table B.2) in the Appendix, the algorithm worked exceptionally well, with an overall accuracy of over 90% for both operating systems.

2.3 Treatments

The recruitment took place on a rolling basis between August and November of 2019. Once a participant registered the first track on the app, he or she 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.[6] 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 varies 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) that the participant individually caused, they were also shown a personalized summary of the previous week.[7] For the remainder of the

[6] 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).

[7] To provide participants with ex-ante information about the congestion costs for particular trips

treatment period, the participants were presented with weekly summaries such that they could observe changes in their external 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 20 % for a buffer to allow for the possibility that some participants had to increase their external costs of transport for idiosyncratic reasons.[8] 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.[9]

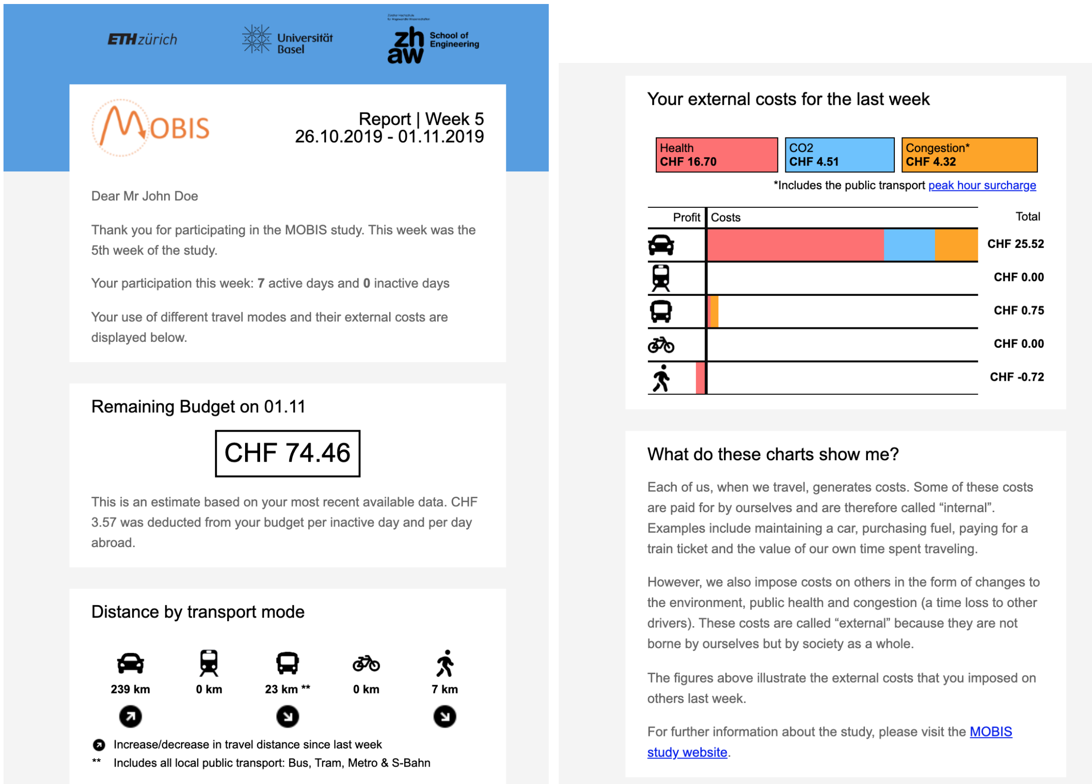
In principle, we could have used any pricing scheme and estimated the participants response to it. We chose the Pigovian rate (i.e., the marginal external costs) for three reasons. First, internalizing the external costs of transport can be motivated and explained to people on normative grounds. The "information only" treatment could thus be interpreted as providing information about societal costs about which the participants were likely not perfectly informed. In contrast, introducing a price unrelated to the external

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.

[8]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.

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

Figure 3: Weekly reports by e-mail



Note: 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.

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.[10] If larger (smaller) 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.

[10]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.

3 The external costs of transport

The health, emissions, noise and congestion costs of the mobility behavior are 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.

3.1 Costs associated with driving

For the calculation of external costs in private road transport, the recorded GPS tracks are 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 are 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.

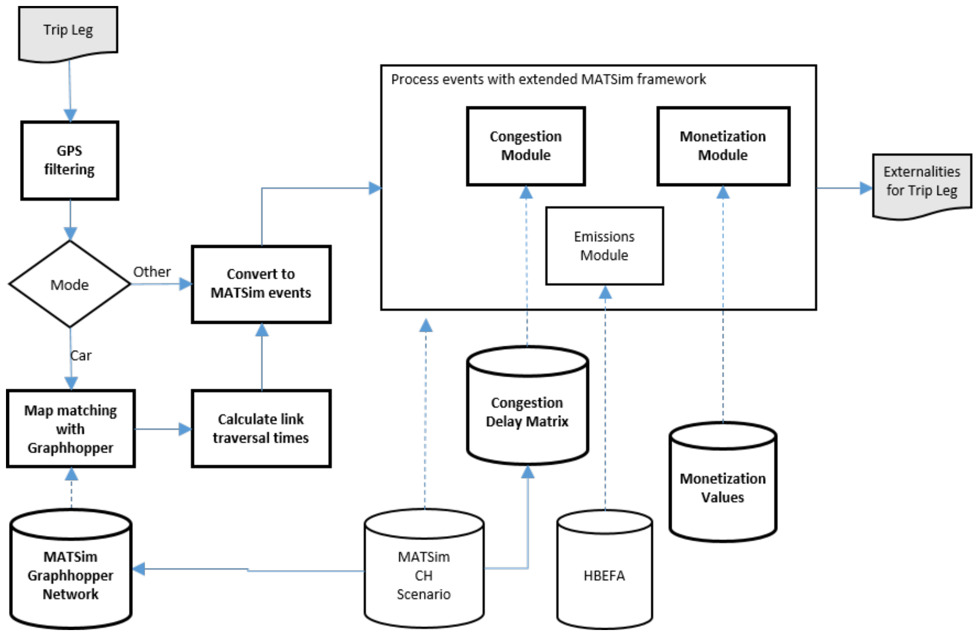
Table 2: Monetization of externalities

Emission	Aspect	Value	Unit
Scenario year		2019	
CO_2	Climate Costs	136.08	CHF/ton
PM_{10} Costs (Healthcare)	Rural	515,497	CHF/ton
	Urban	1,358,461	CHF/ton
NO_x	Regional	7,109	CHF/ton
VTTS	National	25.77	CHF/h [a]

Source: Federal Roads Office - ASTRA (2017), updated for 2019; *a* scaled nominal wage rate

Figure 4 illustrates the externalities pipeline for motorized private transport. 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 4: MATSim-based externalities pipeline



Source: Molloy, Tchervenkov and Axhausen (2021)

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 external costs in the experiment.[11]

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 thus feel 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,

[11]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).

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

Mode	CO_2	PM_{10}	NO_x	Accidents	Noise	Health
Train	0.000066	0.0140	-	0.00066	0.0087	-
Bus	0.0144	0.0437	0.5440	0.0141	0.0257	-
Tram	-	-	-	0.0126	0.0075	-
Bicycle	-	-	-	0.257	-	-0.1870
Walk	-	-	-	0.075	-	-0.1863

Notes: Federal Roads Office - ASTRA (2017) does not report CO_2 values for Tram, but this would be small. Negative costs indicate an external benefit. The external costs of driving vary over time and space and were computed within MATSim (see Fig. 4).

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 a 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 experience 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 variations in working patterns. Municipality pairs were priced if the maximum hourly transit trip count during peak hour was greater than three times the average hourly transit trip count during the daily off-peak (9am - 5pm) for that pair. A municipality could also be paired with itself if the above criteria was met, and the direction of the peak hour flow is not considered. If the trip was partially in both the peak and off-peak periods, only the proportion of the travel duration that overlaps 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 econometrics involved in estimating the average treatment effect (ATE) are straightforward. After aggregating the data to the person-day level, the ATE can be estimated by

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

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 $s \in (1, ..., 56)$. This can be the total quantity of external costs, the external costs along a particular dimension (health, climate and congestion), distance traveled in total or by mode, the mode share, or additional outcomes as specified in the results section.

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 particular 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-fixed effects 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-sample fixed effects account for the possibility that respondents may respond differently to the treatment over time. We find a high variation on the first day (presumably due to technical issues and learning by participants) and on day 29 (the first day of the treatment, and some people may have been slow to check their e-mails). We remove these two days from the analysis. 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. We allow for a correlation of the error within participants.

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. This leads to a fully nonparametric regression in which we compare simple means between treated and control observations. However, because weather information is an important predictor for mobility, especially for leisure activities and for active transport modes, we enrich our tracking data with temperature and precipitation data from MeteoSwiss provided on a 1 x 1 km grid.[12] 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,

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

we define the level of "Heat" and and "Cold" for an observed trip j on day t as follows:

$$Heat_{jt} \equiv \max\{t_{jt}^{max} - 25, 0\} \quad (2)$$

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

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. In addition, we included the daily precipitation (recorded between 6 am and 6 am on the following day). To compute the values per person and day (which is our unit of analysis), we take 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).

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, $\alpha^P - \alpha^I$. This cannot be interpreted as the effect of pricing *per se*, as adding a price could make the information more salient or interact with it in some other manner. 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.

To investigate potential differences of the treatment effect along major socio-economic variables, we interact the DiD terms in eq. (1) 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) 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. We prefer this approach to taking logarithms due to the possible presence of heteroskedasticity, which can lead to a bias in log-linearized regressions, and the presence of zeroes in the data.[13]

[13]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).

5 Results

5.1 Tracking summary

Table 4 shows the summary statistics of the tracking data for overall travel. Table A.1 in the Appendix shows the data separately for each mode.

Table 4: Tracking summary statistics

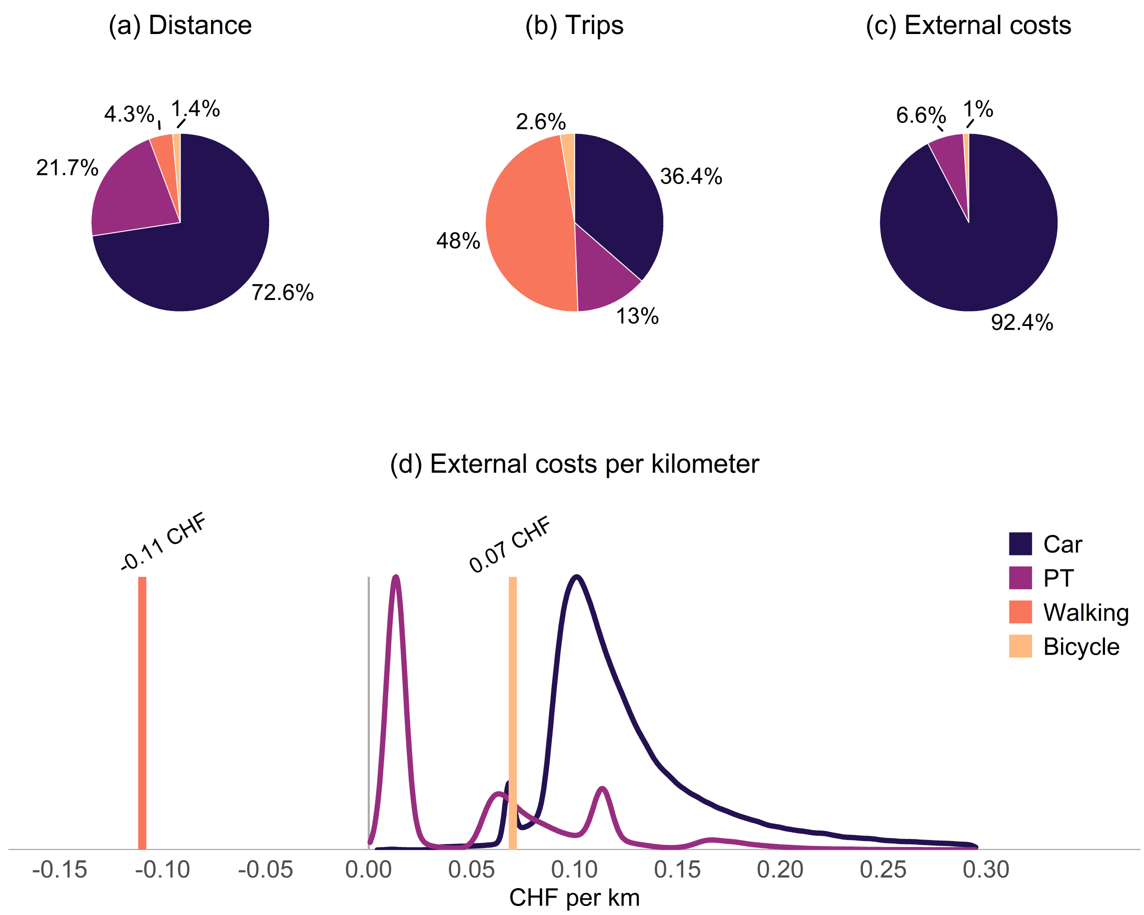
		Pre-treatment			Post-treatment		
Dimension	Outcome	Control	Info	Pricing	Control	Info	Pricing
Ext. costs	Total	4.498	4.579	4.686	4.218	4.252	4.218
(CHF)		(5.690)	(5.650)	(5.811)	(5.392)	(5.549)	(5.408)
	Congestion	1.045	1.076	1.151	0.858	0.878	0.904
		(1.584)	(1.582)	(1.705)	(1.445)	(1.539)	(1.524)
	Climate	0.880	0.883	0.894	0.849	0.838	0.829
		(1.295)	(1.293)	(1.295)	(1.235)	(1.286)	(1.222)
	Health	2.573	2.619	2.640	2.510	2.536	2.485
		(3.551)	(3.546)	(3.609)	(3.476)	(3.561)	(3.460)
Private costs		26.103	26.683	26.827	25.719	25.988	25.476
(CHF)		(33.824)	(34.080)	(34.587)	(33.403)	(34.068)	(33.134)
Tracking	Trips	4.85	4.87	4.90	4.67	4.64	4.71
		(3.11)	(3.08)	(3.04)	(2.89)	(2.88)	(2.92)
	Distance (km)	47.408	48.388	49.854	45.954	47.789	47.784
		(55.789)	(54.978)	(57.720)	(54.331)	(56.246)	(55.269)
	Duration (min)	96.364	97.176	98.453	92.698	95.443	95.421
		(92.890)	(88.568)	(93.527)	(88.397)	(93.972)	(92.736)

Notes:
Average values per participant and day during the experiment (standard dev. in parentheses).

Figure 5 displays the mode distribution recorded in MOBIS in terms of distance (panel a), number of trips (panel b) and external costs (panel c). Almost two-thirds 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.

We observe a seasonal variation in the travel distance by mode, which translates to a negative trend in external costs. Including a control group allows us to control for 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.

Figure 5: 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 public transport vary over time and space.

Before using the data for analysis, it was cleaned using some routine procedures that check for plausibility and remove obviously problematic data. 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 limites 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. Due to the randomization and the presence of the control group, these are the causal effects of introducing price incentives and/or providing information. 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.

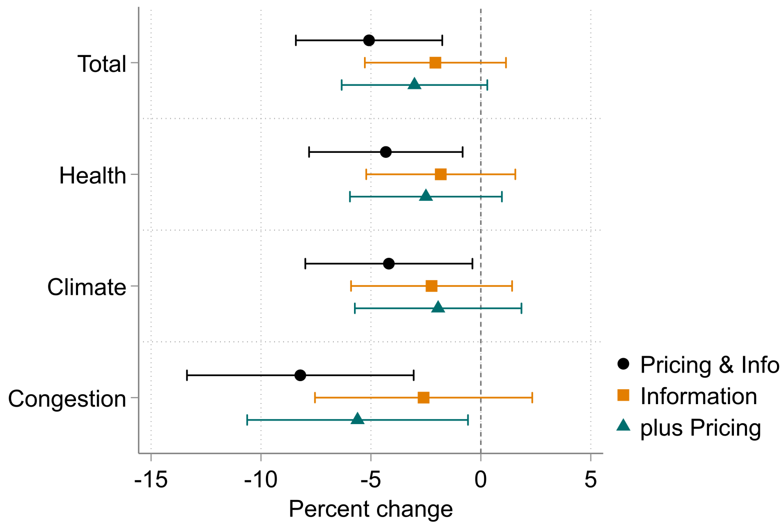
Table 5: Average treatment effects on external costs

	Total Ext. Cost		Health Cost		Climate cost		Congestion Cost	
Pricing	-0.215**	-0.215**	-0.108**	-0.109*	-0.036*	-0.035*	-0.070**	-0.070**
	(0.069)	(0.072)	(0.043)	(0.045)	(0.016)	(0.017)	(0.023)	(0.023)
Information	-0.087	-0.091	-0.046	-0.048	-0.019	-0.020	-0.022	-0.022
	(0.069)	(0.069)	(0.043)	(0.043)	(0.016)	(0.016)	(0.022)	(0.022)
Pricing-Inf.	-0.127'	-0.125'	-0.063	-0.061	-0.017	-0.016	-0.048*	-0.048*
	(0.071)	(0.071)	(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		0.177**		0.148**		0.055**		-0.026**
		(0.018)		(0.012)		(0.004)		(0.004)
Cold		-0.501**		-0.357**		-0.128**		-0.016
		(0.075)		(0.050)		(0.017)		(0.019)
Adj. R^2	0.232	0.234	0.225	0.227	0.221	0.223	0.268	0.269
Clusters	3,656	3,656	3,656	3,656	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208	161,208	161,208	161,208	161,208

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 ("rain") is measured in mm per hour; heat and cold are as defined be eqs. (2)-(3). All regressions include calendar day, day of study and person fixed effects.

Figure 6 shows the proportional reduction of the external costs of transport caused by the two treatments. The proportional effect for the external costs related to health and climate is comparable to the effect on total external costs, whereas the effect for congestion was somewhat stronger. There may be a treatment effect associated with providing only information (yellow bars), but this effect is weaker and not statistically significant at conventional levels, at least not for the overall sample (see below for a sub-sample analysis). The figure also shows the differential effect, which is the causal effect of adding pricing to information. This effect is statistically significant for congestion, indicating that including a monetary component is relatively more important for congestion than for the other types of external costs. A possible explanation is that the external costs in terms of health and climate may be more salient (and more intuitive) than the external costs associated with congestion.

Figure 6: Treatment effect on the external costs of transport



Note: The figure shows the average treatment effect for overall travel from Table 5. Proportional effects were computed by scaling the regression coefficients by the average external cost of the control group during the treatment period. The bars denote 95%-confidence intervals.

An event-study graph of the treatment effect by study day reveals no pronounced learning effects (see Fig. A.2 in the Appendix), but lack of statistical power for a given day leads to very noisy estimates. Furthermore, as part of a series of robustness tests (see below), we find no difference between the ATE as estimated based on weeks 5 & 6 as opposed to weeks 7 & 8. This supports our underlying hypothesis of a stable treatment effect with respect to the treatment time.

Figure A.3 in the Appendix shows the treatment effect for the external costs associated with driving and public transport (the regression estimates underlying this figure are shown in Table A.12). Whereas the proportional reduction of the external costs is similar for driving as for overall travel, we find no significant effect on total costs for PT travel, and a positive effect for health and climate costs. This suggests a mode shift that we will address in more detail below.

On average, the respondents in the pricing group reduced their external costs of transport by 5.1% relative to the control group. To interpret the magnitude of this effect, we have to 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.

To compute the private costs of transport, we make the following assumptions: The private cost of active transport (cycling and walking) is zero; we thus abstract from the purchase or rental price of bicycles. 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 multiplying the half-fare ticket with a factor that is less than one. The level of the factor is determined by comparing the cost of a regional PT subscription with the corresponding cost if one were to buy a daily pass on 22 days per month.[14] 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.

Given these assumptions, the average daily private cost of transport for the control group during the treatment period was CHF 25.72. The external cost was CHF 4.22, which corresponds to a price increase of 16.4%. Dividing the ATE by the average external costs leads to a proportional reduction of 5.1% (this is the point estimate in Figure 6). 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 -5.1/16.4=-0.31.[15] This implies that introducing a transport pricing scheme 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 3.1% in the short run.

Table 6 shows the sensitivity of the results with respect to controlling for the different levels of fixed effects, the weather and using a control group. The preferred model is in column (1). Adding the weather variables has no effect on the ATE, as already shown above. 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; 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 -0.38 if calendar day FE are included, and between -0.57 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 that is exposed to the same shocks as the treatment group. The elasticity is significantly over-estimated

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

[15]This is the elasticity of *external costs* with respect to the *total price* of transport. The elasticity of the external costs with respect to the externality-related price cannot be computed as the latter is zero before the treatment.

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	-0.215**	-0.216**	-0.228**	-0.231**	-0.218**	-0.214**	-0.470**	-0.536**				
	(0.072)	(0.072)	(0.064)	(0.064)	(0.072)	(0.072)	(0.052)	(0.054)				
Information	-0.087	-0.091	-0.101'	-0.105'	-0.088	-0.085*	-0.341**	-0.408**				
	(0.069)	(0.069)	(0.060)	(0.060)	(0.069)	(0.069)	(0.049)	(0.051)				
Precipitation		0.0002		0.002		-0.003**		-0.005*		0.003		-0.006
		(0.004)		(0.004)		(0.002)		(0.002)		(0.007)		(0.004)
Heat		0.177**		0.177**		0.042**		0.024**		0.193**		0.043**
		(0.018)		(0.018)		(0.006)		(0.005)		(0.030)		(0.010)
Cold		-0.501**		-0.499**		-0.250**		-0.189**		-0.388**		-0.256**
		(0.075)		(0.075)		(0.039)		(0.037)		(0.136)		(0.076)
Post									-0.310**	-0.312**	-0.470**	-0.584**
									(0.106)	(0.106)	(0.052)	(0.057)
Prop. effect	-0.051	-0.051	-0.054	-0.055	-0.052	-0.051	-0.111	-0.127	-0.066	-0.067	-0.100	-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	-0.310	-0.312	-0.330	-0.333	-0.315	-0.310	-0.679	-0.774	-0.378	-0.381	-0.574	-0.714
	(0.104)	(0.104)	(0.092)	(0.092)	(0.104)	(0.104)	(0.078)	(0.080)	(0.128)	(0.127)	(0.060)	(0.065)
Person FE	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑
Cal. day FE	☑	☑	☑	☑	☐	☐	☐	☐	☑	☑	☐	☐
Study day FE	☑	☑	☐	☐	☑	☑	☐	☐	☐	☐	☐	☐
Adj. R^2	0.232	0.234	0.232	0.234	0.229	0.230	0.229	0.229	0.228	0.229	0.224	0.225
Cluster	3,656	3,656	3,656	3,656	3,656	3,656	3,656	3,656	1,193	1,193	1,193	1,193
N	161,208	161,208	161,208	161,208	161,208	161,208	161,208	161,208	52,229	52,229	52,229	52,229

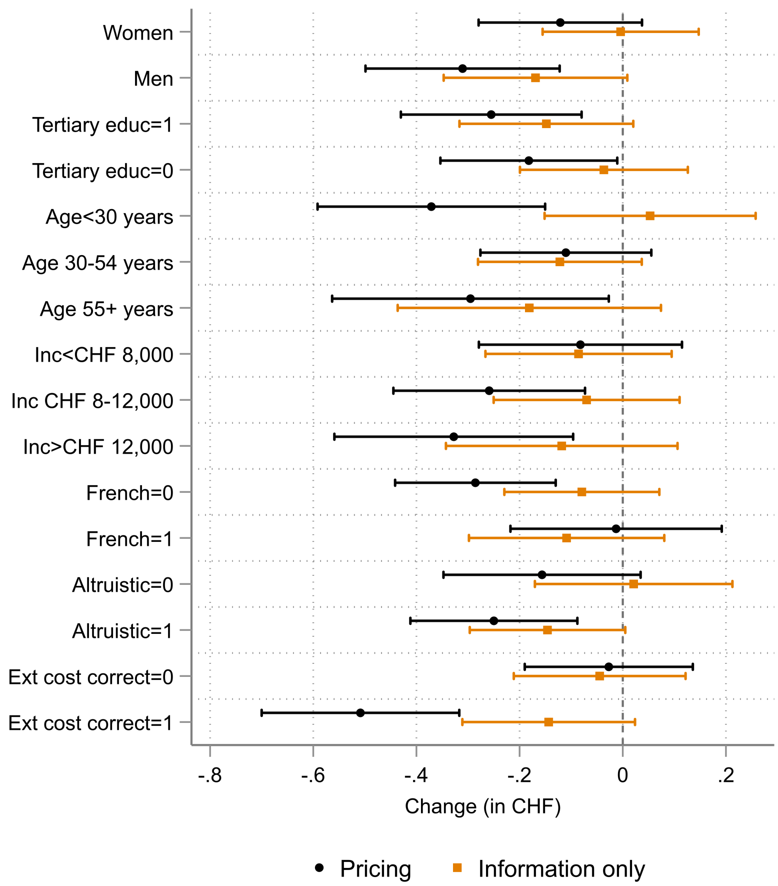
Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors in parentheses and clustered at participant level. The dummy variable "post" takes the value of one during the treatment period (study days 29-56), and zero otherwise. The daily external cost and the price increase implied by Pigovian transport pricing is based on the control group during the treatment phase and is 4.22 CHF and 16.4%, respectively (columns 1-8). For the before-vs.-after analysis in columns (9)-(12), the daily external cost (CHF 4.69) and price increase (17.5%) was computed for the pricing group during the observation phase.

5.3 Effect heterogeneity

The overall effect could mask heterogeneous responses within different segments of the population. Figure 7 shows the treatment effects on total external costs for different subgroups of the sample; the regression results, including those pertaining to the cost categories, are presented in Tables A.3-A.10 in the Appendix. We find that men respond more strongly to both treatments than women, and the difference is particularly pronounced for congestion costs.[16] The treatment effect also differs across the language regions. However, for most socio-demographic variables, we do not find statistically significant differences, or only for cost subgroups.

[16]The graphs show the coefficient on the base category (e.g., for women, whose gender dummmy is coded as male=0) and the sum of the coefficient on the base category and the coefficient on the interaction term (e.g., pricing x male). Whether two groups have statistically different effects can be seen from the p-value of the coefficient of the interaction term, which is reported in Tables A.3-A.10.

Figure 7: Treatment effect heterogeneity



Notes: Results from including interaction dummies in the DiD regression. The bars denote 95%-confidence intervals. The results are given in CHF per day.

In the final survey 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". The study participants that scored above the median in terms of the altruistic index responded significantly not only to pricing, but also to information alone. There were no statistically significant differences along the other three "values" dimensions.

The concept of (and our definition for) the external costs of transport was explained to the participants of the treatment groups at the beginning of the treatment. We included a question in the final survey to gauge the heterogeneity with respect to the uptake of this information. More specifically, the participants in the treatment groups were asked to identify the correct definition of "external costs of transport" among a list of four possible options. The last group of coefficients in Figure 7 shows the ATE separately for those that could correctly identify the definition of the external costs of transport used in the experiment (45% of the sample) vs. those that could not. The results imply that the treatment effect for pricing is driven by the respondents that answered the question correctly. On the other hand, knowledge about the definition of external costs was not generally associated with an increased response to the information-only treatment.

The interaction effects shown in Figure 7 and Tables A.3-A.10 are univariate correlations and have no causal interpretation. In order to investigate the effect heterogeneity further, 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 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 "most important" variables in the sense that they generate the strongest heterogeneity in the ATE.

Figure A.4 in the Appendix 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.5. 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.

Table 7 shows the results of a multivariate regression that includes 15 potentially relevant categorical variables that are interacted with the DiD term in a multi-variate regression. The variables were selected based on their "importance" measure, the results displayed in Fig. 7) and economic theory.[17] We find that when controlling for these dimensions simultaneously, the pricing effect is stronger (i.e., more negative) for participants that (i) answered the "exam question" correctly, (ii) are below 30 years of age and

[17] We are aware that a multivariate regression will not necessarily identify causal effects. However, to the extent that the covariates are correlated with each other, the multivariate regression provides "better controlled" estimates than the univariate regressions that are the basis for Fig. 7.

Table 7: Multivariate interactions

	Total costs		Health Cost		Climate Cost		Congestion Cost	
	Pricing	Info	Pricing	Info	Pricing	Info	Princing	Info
Base	0.258	0.128	0.146	0.050	0.056	0.057	0.056	0.021
	(0.178)	(0.178)	(0.109)	(0.116)	(0.04)	(0.041)	(0.056)	(0.048)
Correct=1	-0.401**	-0.099	-0.253**	-0.071	-0.087**	-0.037	-0.061'	0.01
	(0.109)	(0.102)	(0.068)	(0.064)	(0.025)	(0.023)	(0.035)	(0.032)
Age≥55	-0.112	0.009	-0.079	-0.032	-0.048	-0.001	0.015	0.042
	(0.151)	(0.147)	(0.094)	(0.09)	(0.035)	(0.035)	(0.048)	(0.042)
Age≤30	-0.406**	0.009	-0.239**	0.074	-0.100**	0.024	-0.067'	0.056
	(0.132)	(0.117)	(0.081)	(0.074)	(0.031)	(0.027)	(0.041)	(0.037)
Egoistic=1	0.018	-0.05	0.036	-0.039	0.021	-0.031	-0.039	0.02
	(0.109)	(0.108)	(0.067)	(0.068)	(0.025)	(0.025)	(0.034)	(0.032)
Altruistic=1	-0.138	-0.224*	-0.083	-0.117'	-0.027	-0.045'	-0.027	-0.063*
	(0.118)	(0.110)	(0.073)	(0.071)	(0.027)	(0.025)	(0.038)	(0.032)
Hedonic=1	0.111	-0.065	0.051	0.007	0.018	0.005	0.042	-0.076*
	(0.114)	(0.11)	(0.069)	(0.068)	(0.027)	(0.025)	(0.037)	(0.032)
Biospheric=1	-0.164	0.094	-0.097	0.053	-0.036	0.022	-0.031	0.018
	(0.122)	(0.117)	(0.077)	(0.075)	(0.028)	(0.027)	(0.038)	(0.033)
Income>12k	-0.116	0.021	-0.013	0.028	-0.006	-0.016	-0.097*	0.009
	(0.141)	(0.135)	(0.084)	(0.085)	(0.034)	(0.031)	(0.049)	(0.042)
Income<8k	0.205'	-0.021	0.081	-0.049	0.043	-0.003	0.082*	0.031
	(0.121)	(0.117)	(0.075)	(0.072)	(0.027)	(0.027)	(0.036)	(0.035)
HH size>4	-0.199	0.245	-0.14	0.149	-0.089*	0.046	0.031	0.05
	(0.186)	(0.186)	(0.117)	(0.107)	(0.044)	(0.051)	(0.047)	(0.053)
HH size<3	-0.181	-0.057	-0.04	-0.014	-0.042	-0.031	-0.099**	-0.012
	(0.118)	(0.112)	(0.073)	(0.07)	(0.028)	(0.026)	(0.038)	(0.034)
French=1	0.204'	0.003	0.115	-0.032	0.037	-0.017	0.053	0.052
	(0.120)	(0.109)	(0.077)	(0.069)	(0.027)	(0.026)	(0.036)	(0.033)
Tert. educ	-0.052	-0.056	-0.07	-0.039	-0.011	-0.022	0.03	0.005
	(0.112)	(0.107)	(0.068)	(0.067)	(0.026)	(0.025)	(0.037)	(0.031)
Male=1	-0.195'	-0.165'	-0.066	-0.02	-0.021	-0.022	-0.108**	-0.123**
	(0.109)	(0.100)	(0.068)	(0.062)	(0.025)	(0.023)	(0.033)	(0.031)
Weekend=1	0.12	0.179	0.058	0.089	0.041	0.019	0.022	0.071*
	(0.122)	(0.118)	(0.077)	(0.075)	(0.027)	(0.026)	(0.035)	(0.036)
Adjusted R^2	0.232		0.225		0.221		0.269	
Clusters	3,486		3,486		3,486		3,486	
N	155,517		155,517		155,517		155,517	

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. Standard errors (in parentheses) are clustered at the participant level. The "Pricing" and "Info" 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.

Table 8: 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
Male=1	-5.88	-10.48	-1.27	16.47	16.21	16.73	-0.36	-0.64	-0.08	0.012	80,338
Male=0	-4.09	-8.81	0.63	16.32	16.06	16.58	-0.25	-0.54	0.04	0.089	80,863
Correct=1	-10.57	-14.93	-6.20	16.40	16.22	16.59	-0.64	-0.91	-0.38	<0.001	103,685
Correct=0	0.03	-3.72	3.78	16.40	16.22	16.58	0.00	-0.23	0.23	0.989	115,038
French=1	2.33	-4.72	9.37	16.79	16.40	17.19	0.14	-0.28	0.56	0.518	42,014
French=0	-7.39	-11.20	-3.57	16.29	16.09	16.50	-0.45	-0.69	-0.22	<0.001	119,190
Inc. <8000	-2.51	-51.10	20.04	16.15	15.87	16.44	-0.16	-0.51	0.20	0.392	59,779
Inc >=8000	-6.28	-10.29	-2.28	16.55	16.31	16.79	-0.38	-0.62	-0.14	0.002	101,425

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.

(iii) men, and it is weaker for (iv) French speakers and (v) those with a household income below CHF 8,000 per month. The effect of information alone on total external costs is stronger for (i) altruists and (ii) men. The effect heterogeneity for the individual cost dimensions is overall similar, but there are also some differences. For example, we see that people with a high household income reduced their congestion costs by more than the reference category.

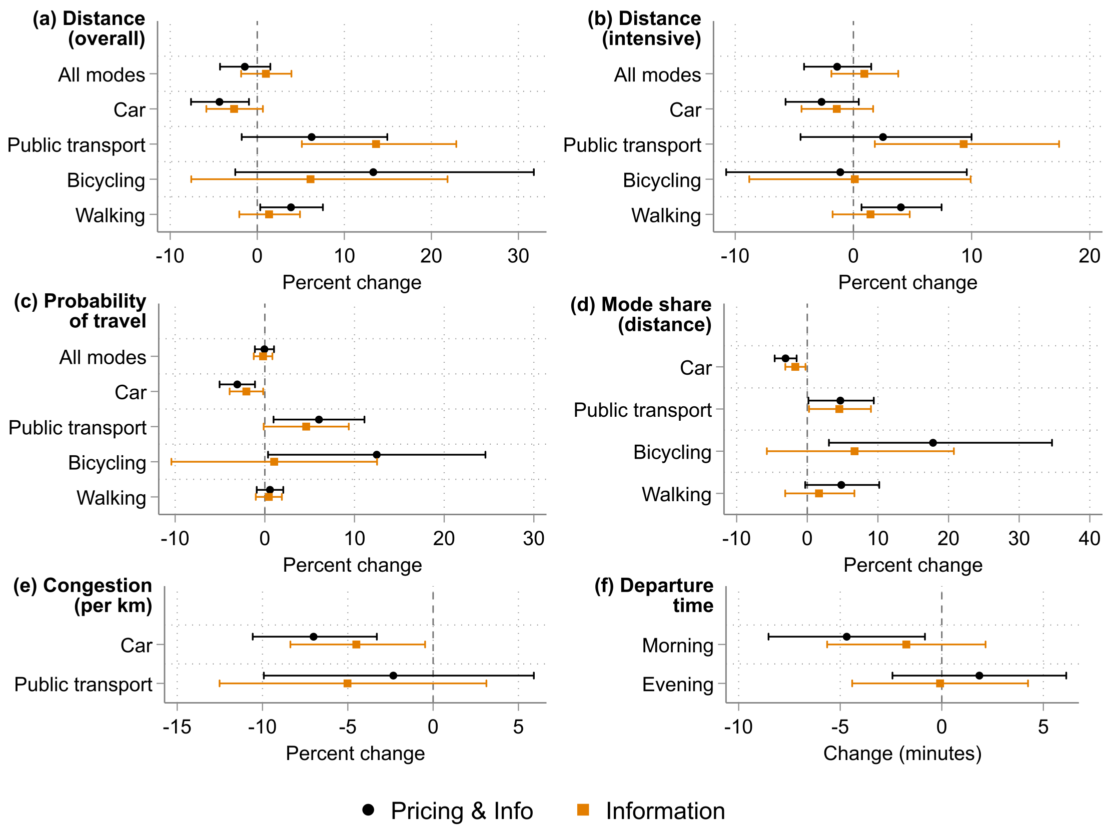
Table 8 presents the proportional effects, price increases and resulting elasticities for the sub-samples for which we found the largest effect heterogeneity of the pricing treatment. The elasticity of the participants that correctly identified the external costs in the "exam" question is in fact -0.64, whereas the elasticity of the rest of the sample (around 55% of the participants in the treatment groups) is precisely centered around zero. The ATE is thus exclusively driven by those participants that understood the nature of the experiment. The table also shows the elasticity for men vs. women, French- vs. German-speakers and people with a monthly household income above/below CHF 8,000.

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 8 shows the effect of the pricing treatment on various outcomes. The regression results underlying this figure are shown in Tables A.11 - A.14 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 even more pronounced 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.

Figure 8: Mechanisms underlying the reduction in external costs



Note: The bars denote 95%-confidence intervals, corresponding to a two-sided test at $p<0.05$. 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 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.

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). There was no reduction in crowding for public transport, nor was there an effect on PT departure time. The treatment did not change the average car and PT speed on a daily level.

5.5 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 introduction survey, we included three questions designed to elicit respondents' preferences about 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 aimed at 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:

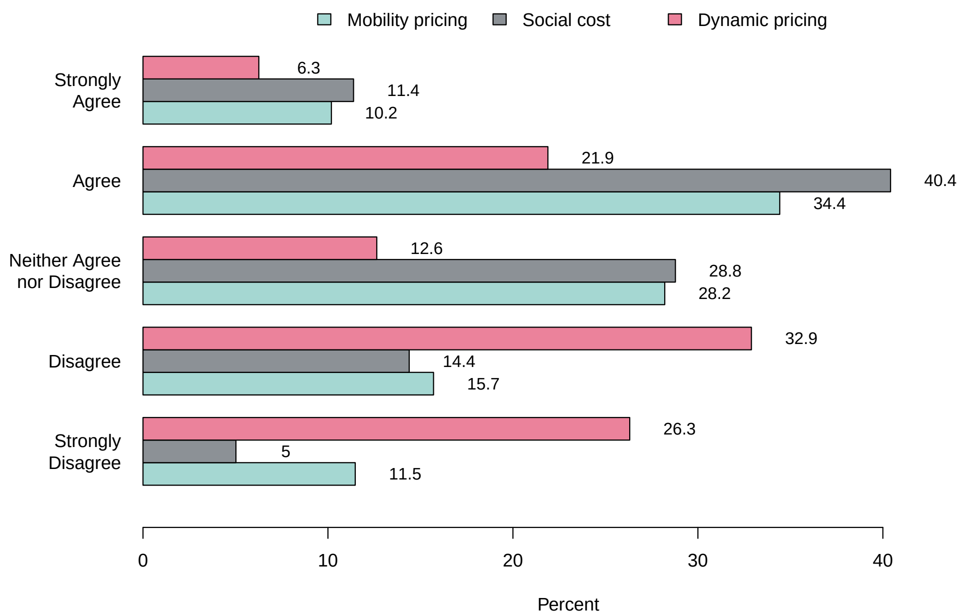
The price for mobility should reflect the social cost (e.g., health, environment, congestion).
The transport network should be used more efficiently by introducing dynamic pricing.

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.[18] 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.

In the final survey, we followed up on the issue of revenue use. The a 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.1). The revenue neutrality embedded in the question about mobility pricing thus likely led to a lower level of approval, *ceteris paribus*.

[18]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



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

5.6 Robustness tests

In this subsection, we address problems that could potentially threaten the validity of our results.

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 9 shows the results from regressing the daily external costs of travel during the observation phase on dummies denoting membership in 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.15), 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.

Table 9: Trends in external costs

	All Travel	All Travel	Car	PT	Bicycle	Walking
Info	8.574	-4.206	-6.704	4.167	-0.452	-1.217
	(12.071)	(15.902)	(16.114)	(2.887)	(0.801)	(0.743)
Pricing	18.853	5.115	-0.062	6.375*	0.236	-1.434*
	(12.245)	(16.079)	(16.268)	(3.039)	(0.936)	(0.721)
Day of study trend	-1.115**	-1.703**	-1.715**	0.099	-0.006	-0.081**
	(0.397)	(0.539)	(0.542)	(0.102)	(0.026)	(0.027)
Pricing x Day-of-study trend		0.921	0.864	0.033	-0.022	0.046
		(0.640)	(0.647)	(0.102)	(0.032)	(0.046)
Info x Day-of-study trend		0.957	0.900	-0.090	0.006	0.040
		(0.638)	(0.642)	(0.104)	(0.030)	(0.040)
Adj. R^2	0.005	0.005	0.005	0.009	0.007	0.012
N	83,081	83,081	83,081	83,081	83,081	83,081

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.

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 over-writing 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 a systematically different 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 10. 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).[19] 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

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

corrected a mode on a given day, and zero otherwise. We find no differential correction on the extensive margin.

Table 10: Mode Correction

	Corrections	Corrections	Probability (Corr.)	Probability (Corr.)
Pricing	1.024	1.009	0.002	0.0003
	(0.066)	(0.065)	(0.006)	(0.007)
Information	0.981	0.976	0.000	-0.001
	(0.063)	(0.063)	(0.007)	(0.007)
Controls	☐	☑	☐	☑
N	74,899	72,759	74,858	72,718

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors (in parentheses) clustered at participant level. The dependent variable in the first two columns is the number of mode corrections per day. The coefficients are proportional effects, estimated using a ppml model. Columns 3-4 display the marginal effects from logit regressions. All regressions control for calendar and study day FE.

To further test the robustness of our results, we re-run 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 11, along with the baseline results. Although the point estimates change somewhat due to the different sample, the effects remain largely stable. This implies that our results are unlikely to be driven by strategic mode correction.

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 11) 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.

Table 11: Subsample analyses

	Baseline	w/o corrections	w/o weeks 7-8	w/o weeks 5-6	w/o zeroes
Pricing	-0.215**	-0.237**	-0.234**	-0.228**	-0.216**
	(0.072)	(0.080)	(0.082)	(0.086)	(0.074)
Information	-0.087	-0.075	-0.080	-0.092	-0.091
	(0.069)	(0.078)	(0.081)	(0.084)	(0.071)
Constant	4.462**	4.711**	4.527**	4.486**	4.664**
	(0.020)	(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	3,656	3,656	3,656	3,656	3,656
Adj.R^2	0.232	0.238	0.235	0.232	0.238
N	161,208	139,703	121,311	122,977	154,520

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.

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 (and our algorithm) impute an uninterrupted activity "at home" that spans the trips on Friday and Monday, then we assume that this person stayed at home and assign a travel distance of zero. However, the imputation of activities 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 11. Due to the exclusion of zero-travel days, the average daily external cost increases, and the proportional effect and the elasticity therefore decrease somewhat, but the results are overall quite stable.

6 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 -0.31. Whereas the information-only treatment had an effect for subgroups of the population (such as altruists), the effect is not statistically significant for the sample overall. 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 gender, age, income, household size and language region, and also with the degree to which participants engaged with the experiment.

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 about 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.

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. Furthermore, 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 are a number of arguments for expecting larger effects in the long run. With a permanent introduction of transport pricing, additional margins of response would 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. Furthermore, the behavioral response was concentrated among those respondents that understood the concept of external costs underlying the experiment. 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.

Our experiment shows that multi-modal transport pricing works in practice. The required technology is available, and a number 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.

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.

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, 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.

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.

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

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, INRIX Research.

Isaksen, Elisabeth T., and Bjorn G. Johansen. 2021. "Congestion pricing, air pollution, and individual-level behavioral responses." Mimeo.

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 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.

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

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.

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`.

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.

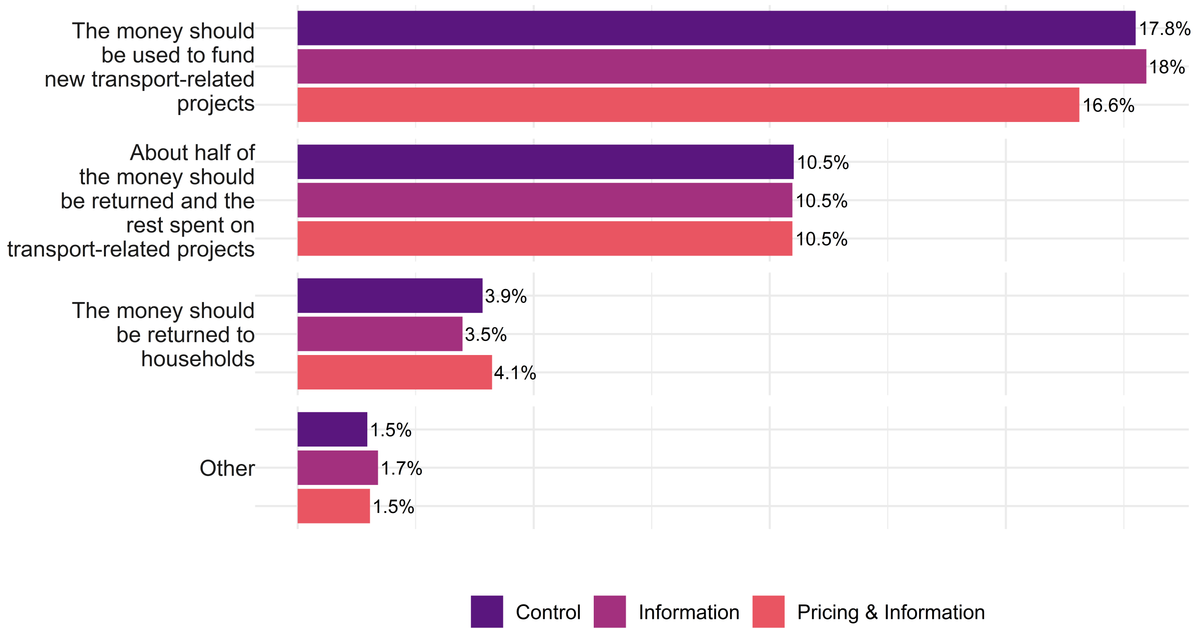
Wu, Linlin, Biao Yang, and Peng Jing. 2016. "Travel mode detection based on GPS raw data collected by smartphones: A systematic review of the existing methodologies." *Information*, 7(4): 67.

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

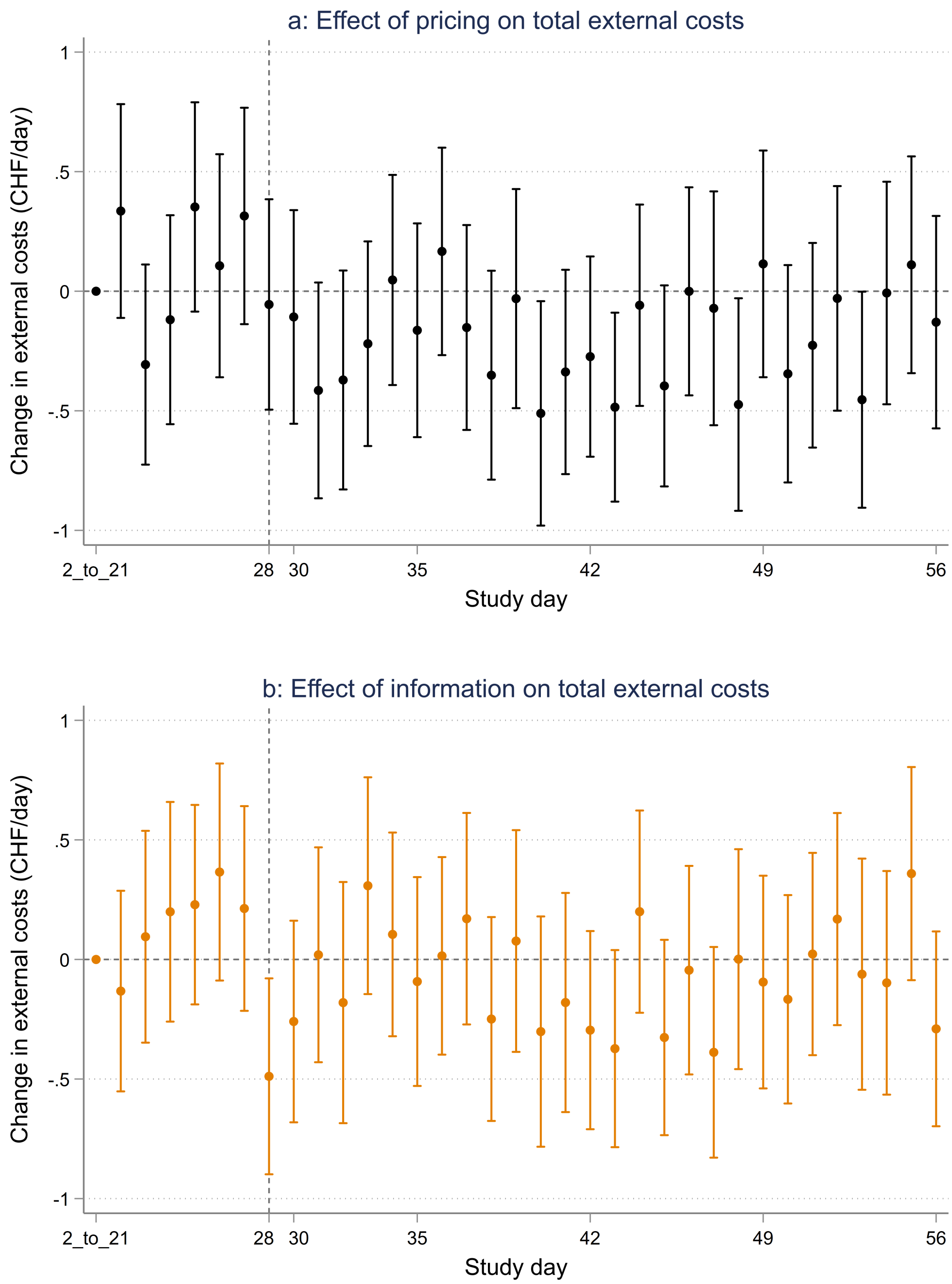
A Additional tables and figures

Figure A.1: How revenue from transport pricing should be used.



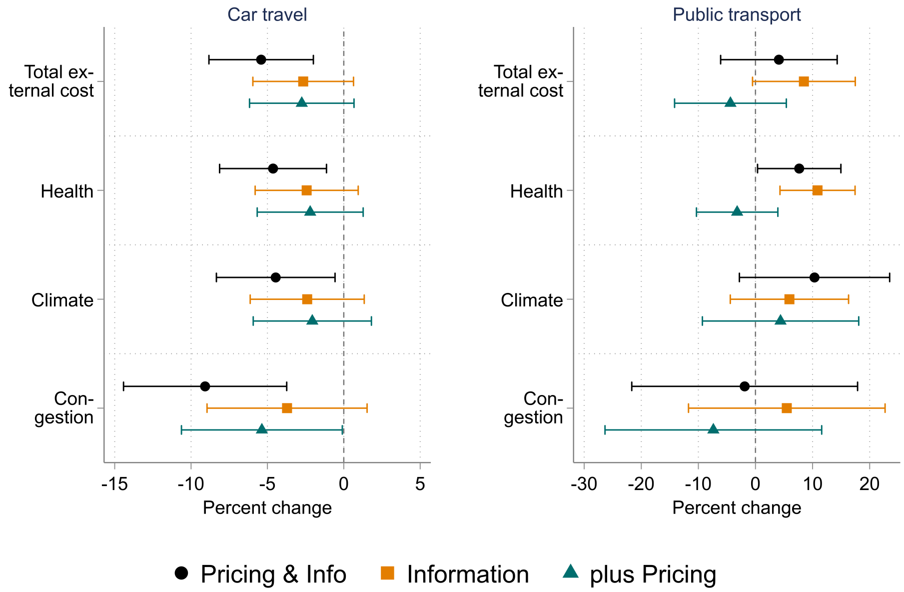
Notes: Based on question: “If dynamic mobility pricing (i.e., prices depending on mode, route and time) were introduced, what should be done with the revenue?”.

Figure A.2: Event study graph



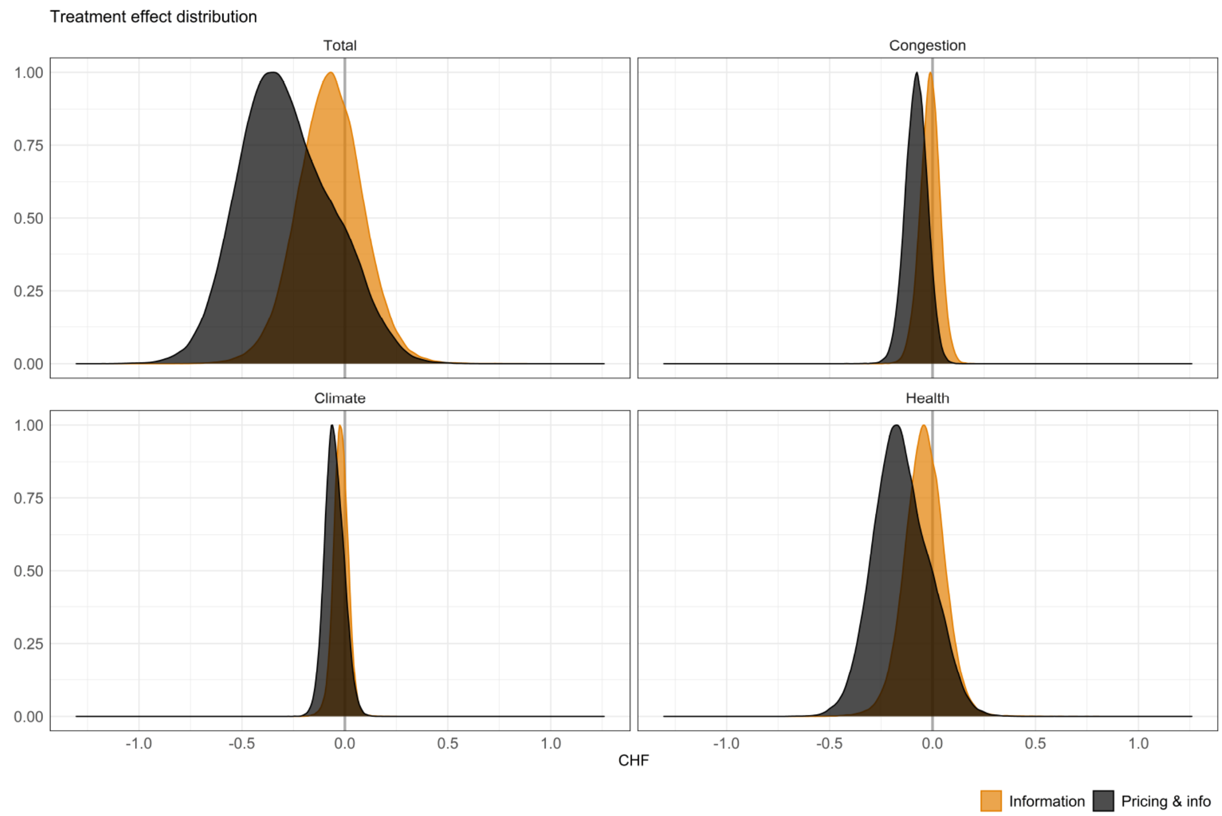
Note: These event study graphs show the effect of pricing plus information (top) and information only (bottom) per study day. These are the coefficients on the interaction terms involving the study day and the treatment dummies. The effects are shown relative to the first three weeks, which is normalized to zero. The bars show 95% confidence intervals. Study day 29 has been removed, making day 30 the first in the treatment period.

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



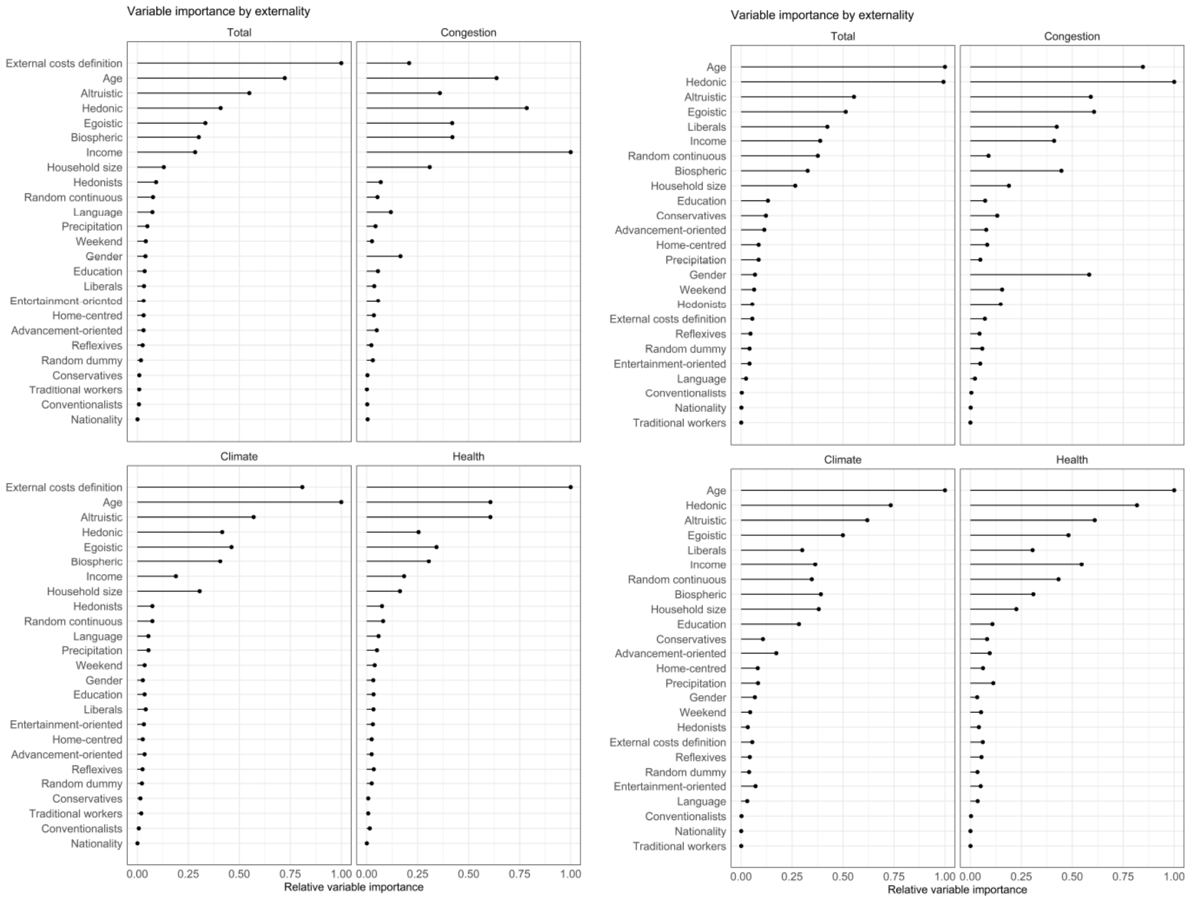
Note: The figure shows the proportional Average Treatment Effect for driving (left) and public transport (right). For additional notes, see Fig. 6 in the main text.

Figure A.4: 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 and the individual cost dimensions considered.

Figure A.5: Variable importance in Causal Forest



Notes: The 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 top left panel shows the conditional treatment effects on total external costs, and the remaining panels show the effects for the dimensions congestion, climate and health.

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	4.39	4.45	4.51	4.12	4.11	4.05	0.28	0.30	0.35	0.27	0.32	0.35
(CHF)		(5.72)	(5.69)	(5.84)	(5.41)	(5.57)	(5.43)	(0.99)	(1.03)	(1.20)	(0.92)	(1.09)	(1.11)
	Congestion	0.94	0.95	0.99	0.75	0.74	0.75	0.11	0.13	0.16	0.10	0.13	0.16
		(1.45)	(1.43)	(1.51)	(1.31)	(1.34)	(1.33)	(0.73)	(0.78)	(0.91)	(0.68)	(0.83)	(0.84)
	Climate	0.86	0.87	0.88	0.83	0.82	0.81	0.02	0.02	0.02	0.02	0.02	0.02
		(1.30)	(1.30)	(1.30)	(1.24)	(1.29)	(1.22)	(0.07)	(0.07)	(0.09)	(0.06)	(0.06)	(0.08)
	Health	2.58	2.63	2.64	2.53	2.55	2.49	0.15	0.16	0.17	0.15	0.17	0.18
		(3.58)	(3.57)	(3.64)	(3.49)	(3.59)	(3.48)	(0.41)	(0.40)	(0.45)	(0.38)	(0.41)	(0.44)
Private cost		24.40	24.99	24.96	24.05	24.20	23.58	1.70	1.69	1.87	1.67	1.79	1.90
(CHF)		(34.01)	(34.32)	(34.81)	(33.58)	(34.29)	(33.36)	(4.53)	(4.29)	(4.85)	(4.42)	(4.50)	(4.67)
Tracking	Distance	34.91	35.76	35.76	34.42	34.70	33.91	9.86	9.98	11.40	9.17	10.65	11.34
	(km)	(48.59)	(49.02)	(49.86)	(48.00)	(49.04)	(47.82)	(32.97)	(31.18)	(35.35)	(30.89)	(33.72)	(33.89)
	Duration	50.62	51.37	51.32	49.94	49.27	48.98	16.32	16.59	17.77	16.50	18.77	18.80
	(min)	(55.87)	(56.99)	(57.94)	(56.54)	(56.89)	(59.43)	(53.26)	(50.28)	(53.61)	(52.68)	(56.90)	(53.51)

		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	0.03	0.04	−0.21	−0.22	−0.22	−0.21	−0.22	−0.22
(CHF)		(0.28)	(0.25)	(0.30)	(0.21)	(0.21)	(0.25)	(0.28)	(0.28)	(0.28)	(0.27)	(0.28)	(0.28)
	Congestion												
	Climate												
	Health	0.05	0.05	0.05	0.04	0.03	0.04	−0.21	−0.22	−0.22	−0.21	−0.22	−0.22
		(0.28)	(0.25)	(0.30)	(0.21)	(0.21)	(0.25)	(0.28)	(0.28)	(0.28)	(0.27)	(0.28)	(0.28)
Private cost													
(CHF)													
Tracking	Distance	0.72	0.67	0.71	0.51	0.50	0.56	1.92	1.98	1.99	1.86	1.94	1.98
	(km)	(3.98)	(3.62)	(4.34)	(3.06)	(2.99)	(3.57)	(2.51)	(2.55)	(2.51)	(2.39)	(2.48)	(2.48)
	Duration	2.49	2.33	2.34	1.86	1.72	1.90	26.93	26.89	27.03	24.40	25.68	25.74
	(min)	(13.70)	(12.03)	(13.58)	(12.79)	(9.72)	(11.14)	(55.05)	(49.89)	(52.77)	(46.82)	(53.48)	(50.94)

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

Table A.2: Effect for car and public transport

(a) Car

	Total External Cost		Health Costs	Climate Costs	Congestion Costs
Pricing	-0.223**	-0.224**	-0.117**	-0.037*	-0.069**
	(0.072)	(0.074)	(0.045)	(0.017)	(0.021)
Information	-0.110	-0.113	-0.062	-0.020	-0.028
	(0.069)	(0.069)	(0.044)	(0.016)	(0.020)
Difference	-0.113	-0.111	-0.056	-0.017	-0.040*
	(0.072)	(0.072)	(0.045)	(0.016)	(0.020)
Precipitation		0.003			
		(0.004)			
Heat		0.186**			
		(0.018)			
Cold		-0.503**			
		(0.076)			
Adj. R^2	0.238	0.239	0.229	0.223	0.274
Clusters	3,656	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208	161,208

(b) Public Transport

	Total External Cost		Health Cost	Climate Cost	Congestion Cost
Pricing	0.011	0.011	0.011*	0.002	-0.002
	(0.014)	(0.014)	(0.005)	(0.001)	(0.010)
Information	0.023'	0.023'	0.016**	0.001	0.006
	(0.012)	(0.012)	(0.005)	(0.001)	(0.009)
Difference	-0.012	-0.012	-0.005	0.001	-0.008
	(0.013)	(0.013)	(0.005)	(0.001)	(0.010)
Precipitation		-0.002*			
		(0.001)			
Heat		-0.006'			
		(0.003)			
Cold		-0.007			
		(0.015)			
Adj. R^2	0.286	0.286	0.235	0.137	0.274
Cluster	3,656	3,656	3,656	3,656	3,656
N	168,208	168,208	168,208	168,208	168,208

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: Effect heterogeneity: Gender and Education

(a) Gender

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.121	-0.079	-0.028	-0.014
	(0.081)	(0.051)	(0.019)	(0.026)
Information	-0.004	-0.035	-0.004	0.035
	(0.077)	(0.049)	(0.018)	(0.024)
Pricing x male	-0.189'	-0.059	-0.016	-0.114**
	(0.104)	(0.064)	(0.024)	(0.032)
Inf. x male	-0.165'	-0.022	-0.029	-0.114**
	(0.097)	(0.061)	(0.022)	(0.029)
Pricing + Pr. x male	-0.311**	-0.138*	-0.044*	-0.128**
	(0.096)	(0.059)	(0.022)	(0.030)
Inf. + Inf. x male	-0.169'	-0.057	-0.033	-0.079*
	(0.091)	(0.057)	(0.021)	(0.028)
Pricing - Inf.	-0.117	-0.045	-0.023	-0.049'
	(0.088)	(0.055)	(0.020)	(0.027)
(Pricing + Pr. x male) - (Inf. + Inf. x male)	-0.141	-0.082	-0.010	-0.049
	(0.112)	(0.069)	(0.026)	(0.035)
Adj. R^2	0.232	0.225	0.221	0.269
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

(b) Education

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.182*	-0.080	-0.031	-0.071**
	(0.087)	(0.054)	(0.022)	(0.027)
Information	-0.037	-0.018	-0.001	-0.017
	(0.083)	(0.052)	(0.019)	(0.026)
Pricing x tertiary educ.	-0.073	-0.063	-0.010	0.000
	(0.104)	(0.064)	(0.024)	(0.033)
Inf. x tertiary educ.	-0.012	-0.060	-0.041'	0.011
	(0.97)	(0.061)	(0.022)	(0.030)
Pricing + Pr. x tertiary educ.	-0.255**	-0.143*	-0.041*	-0.070*
	(0.089)	(0.056)	(0.020)	(0.029)
Inf. + Inf. x tertiary educ.	-0.148'	-0.079	-0.041*	-0.028
	(0.086)	(0.054)	(0.020)	(0.027)
Pricing - Inf.	-0.146	-0.062	-0.030	-0.053'
	(0.099)	(0.061)	(0.023)	(0.030)
(Pricing + Pr. x tertiary educ.) - (Inf. + Inf. x tertiary educ.)	-0.107	-0.065	0.000	-0.042
	(0.103)	(0.064)	(0.023)	(0.033)
Constant	4.462**	2.587**	0.887**	1.003**
	(0.020)	(0.012)	(0.004)	(0.006)
Adj. R^2	0.221	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. The dependent variable contains the external costs of transport aggregated to the person-day level (in CHF). Standard errors (in parentheses) are clustered at the participant level. The dummy variable "male" is equal to one for men. The dummy "tertiary educ." is equal to one for those that have a tertiary education, and zero otherwise. All regressions include individual, calendar day and day of study FE.

Table A.4: Effect heterogeneity: Age

	Total ext. costs	Health costs	Climate costs	Congestion costs
Pricing	-0.110'	-0.046	-0.004	-0.060*
	(0.084)	(0.053)	(0.019)	(0.026)
Information	-0.122	-0.055	-0.026	-0.041
	(0.081)	(0.051)	(0.018)	(0.025)
Pricing x old	-0.185	-0.096	-0.063'	0.026
	(0.145)	(0.086)	(0.033)	(0.049)
Inf. x old	-0.059	-0.066	-0.014	0.020
	(0.137)	(0.084)	(0.032)	(0.038)
Pricing x young	-0.261	-0.166*	-0.075**	-0.020
	(0.122)	(0.076)	(0.028)	(0.037)
Inf. x young	0.175	0.082	0.035	0.058
	(0.112)	(0.071)	(0.026)	(0.036)
Pricing + Pr. x old	-0.295*	-0.142'	-0.067*	-0.087'
	(0.137)	(0.081)	(0.031)	(0.047)
Inf. + Inf. x old	-0.181	-0.120	-0.039	-0.022
	(0.130)	(0.080)	(0.030)	(0.036)
Pricing + Pricing x young	-0.371**	-0.212**	-0.079**	-0.080*
	(0.113)	(0.070)	(0.026)	(0.035)
Inf. + Inf. x young	0.053	0.027	0.010	0.016
	(0.104)	(0.066)	(0.024)	(0.034)
Pricing - Inf.	0.112	0.008	0.022	-0.019
	(0.094)	(0.059)	(0.022)	(0.029)
(Pricing + Pr. x old) - (Inf. + Inf. x old)	-0.114	-0.022	-0.027	-0.051
	(0.175)	(0.105)	(0.041)	(0.0555)
(Pricing + Pr. x young) - (Inf. + Inf. x young)	-0.424**	-0.239**	-0.088**	-0.096*
	(0.137)	(0.085)	(0.032)	(0.043)
Adj. R^2	0.232	0.225	0.221	0.26
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. The dependent variable contains the external costs of transport aggregated to the person-day level (in CHF). Standard errors (in parentheses) are clustered at the participant level. The reference category is the age group 30-54. The dummy "young" is equal to one for people less than 30 years, and the dummy "old" is equal to one for people that are 50 or older (and zero otherwise). All regressions include individual, calendar day and day of study FE.

Table A.5: Effect heterogeneity: Income

	Total ext. costs	Health costs	Climate costs	Congestion costs
Pricing	-0.259**	-0.135*	-0.049*	-0.076*
	(0.095)	(0.059)	(0.021)	(0.031)
Information	-0.070	-0.025	-0.009	-0.036
	(0.069)	(0.058)	(0.021)	(0.029)
Pricing x high inc.	-0.069	0.000	0.002	-0.071
	(0.135)	(0.081)	(0.032)	(0.045)
Inf. x high inc.	-0.048	-0.011	-0.035	0.002
	(0.129)	(0.081)	(0.030)	(0.040)
Pricing x low inc.	0.177	0.075	0.036	0.066'
	(0.119)	(0.075)	(0.027)	(0.035)
Inf. x low inc.	-0.016	-0.047	-0.006	0.037
	(0.110)	(0.069)	(0.025)	(0.033)
Pricing + Pr. x high inc.	-0.328**	-0.134'	-0.046	-0.147**
	(0.118)	(0.071)	(0.029)	(0.040)
Inf. + Inf. x high inc	-0.118	-0.036	-0.044	-0.038
	(0.115)	(0.072)	(0.027)	(0.036)
Pricing + Pr. x low inc.	-0.082	-0.059	-0.013	-0.010
	(0.100)	(0.064)	(0.023)	(0.028)
Inf. + Inf. x low inc	-0.086	-0.073	-0.015	0.001
	(0.092)	(0.057)	(0.021)	(0.028)
Pricing - Inf.	-0.189'	-0.109	-0.040	-0.040
	(0.112)	(0.070)	(0.025)	(0.036)
(Pricing + Pr. x high inc.) - (Inf. + Inf. x high inc.)	-0.209	-0.098	-0.002	-0.109*
	(0.149)	(0.091)	(0.036)	(0.049)
(Pricing + Pr. x low inc.) - (Inf. + Inf. x low inc.)	0.004	0.013	0.002	-0.011
	(0.117)	(0.073)	(0.027)	(0.032)
Adj. R^2	0.232	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. The dependent variable contains the external costs of transport aggregated to the person-day level (in CHF). Standard errors (in parentheses) are clustered at the participant level. The reference category is given by a household income between CHF 8,000 and 12,000 per month. "high inc." and "low inc." are dummy variables denoting respondents that reported a monthly household income of above CHF 12,000 and below CHF 8,000, respectively. All regressions include individual, calendar day and day of study FE.

Table A.6: Effect heterogeneity: Household size

	Total ext. costs	Health costs	Climate costs	Congestion costs
Pricing	-0.135	-0.084	-0.011	-0.041
	(0.086)	(0.053)	(0.019)	(0.027)
Information	-0.082	-0.051	-0.012	-0.019
	(0.084)	(0.053)	(0.019)	(0.026)
Pricing x large size	-0.293	-0.188	-0.108*	0.002
	(0.182)	(0.114)	(0.042)	(0.046)
Inf. x large size	0.265	0.165	0.049	0.051
	(0.182)	(0.106)	(0.049)	(0.051)
Pricing x small size	-0.125	-0.015	-0.035	-0.074*
	(0.111)	(0.068)	(0.026)	(0.036)
Inf. x small size	-0.065	-0.020	-0.027	-0.018
	(0.102)	(0.064)	(0.023)	(0.032)
Pricing + Pr. x large size	-0.429*	-0.272*	-0.119**	-0.038
	(0.175)	(0.110)	(0.041)	(0.044)
Inf. + Inf. x large size	0.183	0.114	0.037	0.032
	(0.176)	(0.102)	(0.048)	(0.049)
Pricing + Pr. x small size	-0.260**	-0.099	-0.046*	-0.114**
	(0.099)	(0.064)	(0.023)	(0.032)
Inf. + Inf. x small size	-0.147	-0.071	-0.039’	-0.037
	(0.090)	(0.059)	(0.021)	(0.028)
Pricing - Inf.	-0.053	-0.033	0.001	-0.021
	(0.098)	(0.061)	(0.022)	(0.030)
(Pricing + Pr. x large size) - (Inf. + Inf. x large size)	-0.612*	-0.386**	-0.156*	-0.070
	(0.238)	(0.144)	(0.061)	(0.062)
(Pricing + Pr. x small size) - (Inf. + Inf. x small size)	-0.113	-0.028	-0.007	-0.078*
	(0.115)	(0.071)	(0.027)	(0.036)
Adj. R^2	0.232	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

Notes: **: $p < 0.01$, *: $p < 0.05$, ’: $p < 0.1$. The dependent variable contains the external costs of transport aggregated to the person-day level (in CHF). Standard errors (in parentheses) are clustered at the participant level. The reference category is a household size with 3 or 4 persons. Participants living in households with fewer than 3 persons are denoted by the dummy “small size”, and those living in households with 5 or more persons are denoted by “large size”. All regressions include individual, calendar day and day of study FE.

Table A.7: Heterogeneity: Attention to experiment

(a) Awareness of experimental rules

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.204'	-0.083	-0.031	-0.089*
	(0.111)	(0.067)	(0.025)	(0.037)
Information	0.003	0.015	0.007	-0.019
	(0.097)	(0.060)	(0.022)	(0.030)
Pricing x aware	-0.016	-0.037	-0.006	0.027
	(0.117)	(0.070)	(0.027)	(0.038)
Inf. x aware	-0.149	-0.101	-0.043'	-0.006
	(0.102)	(0.064)	(0.023)	(0.031)
Pricing + Pr. x aware	-0.220**	-0.120**	-0.037*	-0.062**
	(0.078)	(0.049)	(0.018)	(0.024)
Information + Inf. x aware	-0.146'	-0.085'	-0.036*	-0.025
	(0.077)	(0.048)	(0.018)	(0.024)
Constant	4.462**	2.587**	0.872**	1.003**
	(0.020)	(0.012)	(0.004)	(0.006)
Adj. R^2	0.232	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

(b) Correct answer to externalities question

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.027	0.006	0.005	-0.038
	(0.083)	(0.052)	(0.019)	(0.026)
Information	-0.045	-0.018	-0.003	-0.024
	(0.085)	(0.053)	(0.020)	(0.026)
Pricing x correct	-0.481**	-0.296**	-0.103**	-0.083*
	(0.107)	(0.066)	(0.025)	(0.034)
Inf. x correct	-0.099	-0.060	-0.036	0.003
	(0.098)	(0.061)	(0.022)	(0.030)
Pricing + Pr. x correct	-0.509**	-0.289**	-0.098**	-0.121**
	(0.098)	(0.060)	(0.023)	(0.031)
Inf. + Inf. x correct	-0.144'	-0.077	-0.039*	-0.027
	(0.085)	(0.054)	(0.020)	(0.027)
Adj. R^2	0.232	0.224	0.221	0.268
Clusters	3,486	3,486	3,486	3,486
N	155,517	155,517	155,517	155,517

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. The dependent variable contains the external costs of transport aggregated to the person-day level (in CHF). Standard errors (in parentheses) are clustered at the participant level. The dummy "aware" is equal to one for those participants that reported in the final survey that they were aware that they could make money or reduce their external costs of travel by changing their travel patterns. The dummy "correct" is equal to one for those participants that correctly selected the definition of the external costs of transport among four presented options, and zero otherwise.

Table A.8: Heterogeneity: Intrinsic values (1/2)

(a) "Egoistic" score

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.237*	-0.136*	-0.050*	-0.050
	(0.098)	(0.060)	(0.023)	(0.031)
Information	-0.048	-0.027	-0.001	-0.022
	(0.096)	(0.061)	(0.022)	(0.029)
Pricing x egoistic	0.035	0.044	0.023	-0.032
	(0.107)	(0.066)	(0.025)	(0.034)
Inf. x egoistic	-0.060	-0.029	-0.030	-0.001
	(0.102)	(0.064)	(0.023)	(0.031)
Pricing + Pr. x egoistic	-0.202*	-0.092'	-0.027'	-0.082**
	(0.083)	(0.052)	(0.019)	(0.026)
Information + Inf. x egoistic	-0.108	-0.056	-0.029'	-0.023
	(0.078)	(0.049)	(0.018)	(0.024)
Pricing - Inf.	-0.189	-0.110	-0.051'	-0.028
	(0.118)	(0.073)	(0.027)	(0.036)
(Pricing + Pr. x egoistic) - (Inf. + Inf. x egoistic)	-0.094	-0.037	0.002	-0.060*
	(0.090)	(0.056)	(0.021)	(0.028)
Constant	4.462**	2.587**	0.872**	1.003**
	(0.020)	(0.012)	(0.005)	(0.006)
Adj. R^2	0.221	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

(b) "Hedonic" score

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.292**	-0.150*	-0.050*	-0.093**
	(0.1041)	(0.061)	(0.023)	(0.033)
Information	-0.044	-0.041	-0.017	0.013
	(0.094)	(0.060)	(0.021)	(0.028)
Pricing x hedonic	0.128	0.068	0.023	0.037
	(0.109)	(0.068)	(0.025)	(0.035)
Inf. x hedonic	-0.069	-0.008	-0.004	-0.057'
	(0.101)	(0.064)	(0.023)	(0.030)
Pricing + Pr. x hedonic	-0.164*	-0.081	-0.026	-0.056*
	(0.081)	(0.019)	(0.064)	(0.024)
Inf. + Inf. x hedonic	-0.113	-0.049	-0.020	-0.044'
	(0.078)	(0.049)	(0.018)	(0.025)
Pricing - Inf.	-0.248*	-0.109	-0.033	-0.106**
	(0.119)	(0.073)	(0.027)	(0.038)
(Pricing + Pr. x hedonic) - (Inf. + Inf. x hedonic)	-0.051	-0.032	-0.006	-0.012
	(0.088)	(0.055)	(0.021)	(0.027)
Adj. R^2	0.232	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. The dependent variable contains the external costs of transport aggregated to the person-day level (in CHF). Standard errors (in parentheses) are clustered at participant level. Based on a battery of 16 values-related questions, participants were given an index among the dimensions "altruistic", "hedonic", "egoistic" and "biospheric". The dummies "egoistic" and "hedonic" denote participants with an above-median value for these two dimensions.

Table A.9: Heterogeneity: Intrinsic values (2/2)

(a) “Altruistic” score

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.157	-0.065	-0.020	-0.072*
	(0.097)	(0.061)	(0.022)	(0.031)
Information	0.021	0.018	-0.001	0.002
	(0.098)	(0.063)	(0.022)	(0.028)
Pricing x altruistic	-0.093	-0.070	-0.025	0.002
	(0.107)	(0.066)	(0.025)	(0.034)
Inf. x altruistic	-0.167	-0.098	-0.031	-0.038
	(0.103)	(0.065)	(0.024)	(0.030)
Pricing + Pr. x altruistic	-0.250**	-0.135**	-0.045*	-0.070**
	(0.083)	(0.051)	(0.019)	(0.026)
Information + Inf. x altruistic	-0.146’	-0.080’	-0.030’	-0.036
	(0.077)	(0.048)	(0.018)	(0.025)
Pricing - Inf.	-0.178	-0.083	-0.021	-0.074*
	(0.119)	(0.075)	(0.027)	(0.036)
(Pricing + Pr. x altruistic) - (Inf. + Inf. x altruistic)	-0.104	-0.055	-0.015	-0.034
	(0.089)	(0.055)	(0.021)	(0.028)
Constant	4.462**	2.587**	0.872**	1.003**
	(0.020)	(0.012)	(0.004)	(0.006)
Adj. R^2	0.232	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

(b) “Biospheric” score

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.179’	-0.077’	-0.023	-0.078**
	(0.099)	(0.063)	(0.023)	(0.030)
Information	-0.0084	-0.042	-0.022	-0.020
	(0.105)	(0.067)	(0.024)	(0.030)
Pricing x bio	-0.057	-0.050	-0.020	0.012
	(0.108)	(0.067)	(0.025)	(0.033)
Inf. x bio	-0.005	-0.005	-0.043*	-0.004
	(0.109)	(0.068)	(0.019)	(0.032)
Pricing + Pr. x bio	-0.236**	-0.127*	-0.018	-0.0566*
	(0.082)	(0.050)	(0.019)	(0.026)
Information + Inf. x bio	-0.089	-0.047	-0.021	-0.024
	(0.075)	(0.047)	(0.017)	(0.024)
Pricing - Inf.	-0.095	-0.035	-0.002	-0.059
	(0.126)	(0.080)	(0.029)	(0.036)
(Pricing + Pr. x bio) - (Inf. + Inf. x Bio)	-0.147’	-0.080	-0.025	-0.042
	(0.087)	(0.053)	(0.020)	(0.028)
Adj. R^2	0.232	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

Notes: **: $p < 0.01$, *: $p < 0.05$, ’: $p < 0.1$. The dependent variable contains the external costs of transport aggregated to the person-day level (in CHF). Standard errors (in parentheses) are clustered at the participant level. Based on a battery of 16 values-related questions, participants were given an index among the dimensions “altruistic”, “hedonic”, “egoistic” and “biospheric”. The dummies ”altruistic” and ”biospheric” denote participants with an above-median value for these dimensions. All regressions include individual, calendar day and day of study FE.

Table A.10: Heterogeneity: Language and day of week

(a) Language

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.286**	-0.145**	-0.048**	-0.093**
	(0.079)	(0.049)	(0.018)	(0.025)
Information	-0.079	-0.033	-0.015	-0.031
	(0.077)	(0.048)	(0.018)	(0.024)
Pricing x French	0.273*	0.142*	0.046'	0.085*
	(0.111)	(0.070)	(0.025)	(0.035)
Inf. x French	-0.030	-0.047	-0.014	0.032
	(0.102)	(0.066)	(0.024)	(0.030)
Pricing + Pr. x French	-0.013	-0.001	-0.003	-0.008
	(0.104)	(0.067)	(0.023)	(0.033)
Information + Inf. x French	-0.109	-0.082	-0.029	0.001
	(0.097)	(0.063)	(0.022)	(0.028)
Pricing - Inf.	-0.206*	-0.112*	-0.032	-0.062*
	(0.086)	(0.052)	(0.020)	(0.027)
(Pricing - Pr. x French) - (Inf. - Inf. x French)	0.096	0.076	0.028	-0.009
	(0.124)	(0.081)	(0.028)	(0.038)
Constant	4.462**	2.587**	0.882**	1.003**
	(0.020)	(0.012)	(0.004)	(0.006)
Adj. R^2	0.232	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

(b) Weekend

	Total Ext. Cost	Health Cost	Climate Cost	Congestion Cost
Pricing	-0.251**	-0.126**	-0.047**	-0.078**
	(0.077)	(0.048)	(0.018)	(0.024)
Information	-0.137'	-0.070	-0.024	-0.043'
	(0.075)	(0.047)	(0.017)	(0.024)
Pricing x weekend	0.125	0.059	0.040	0.026
	(0.120)	(0.076)	(0.026)	(0.034)
Inf. x weekend	0.172	0.083	0.017	0.071*
	(0.116)	(0.073)	(0.025)	(0.036)
Pricing + Pr. x weekend	-0.126	-0.067	-0.007	-0.052
	(0.115)	(0.073)	(0.026)	(0.034)
Inf + Inf x weekend	0.034	-0.013	-0.007	0.028
	(0.111)	(0.070)	(0.024)	(0.034)
Pricing - Inf.	-0.113	-0.055	-0.023	-0.035
	(0.078)	(0.049)	(0.018)	(0.024)
(Pricing + Pr. x weekend) - (Inf. + Inf. x weekend)	-0.160	-0.080	0.000	-0.080*
	(0.133)	(0.084)	(0.029)	(0.039)
Adj R^2	0.222	0.225	0.221	0.268
Clusters	3,656	3,656	3,656	3,656
N	161,208	161,208	161,208	161,208

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. The dependent variable contains the external costs of transport aggregated to the person-day level (in CHF). Standard errors (in parentheses) are clustered at the participant level. The reference category are German speakers or those that chose English as their preferred language. The dummy variable "French" is equal to one for French speakers. The weekend dummy is equal to one for week-ends and holidays, and zero otherwise. All regressions include individual, calendar day and day of study FE.

Table A.11: ATE on travel distance

(a) Overall margin

	Distance Total	Distance Car	Distance Public Transport	Distance Bicycle	Distance Walk
Pricing	0.986	0.957*	1.062	1.133	1.039*
	(0.015)	(0.017)	(3)	(0.087)	(0.018)
Information	1.010	0.973	1.136**	1.061	1.014
	(0.015)	(0.017)	(0.045)	(0.075)	(0.018)
Difference	0.976’	0.983	0.935’	1.068	1.025
	(0.014)	(0.017)	(0.034)	(0.084)	(0.018)
Precipitation	0.998*	1.000	0.992**	0.978**	0.994
	(0.001)	(0.001)	(0.002)	(0.006)	(0.001)
Heat	1.046**	1.059**	1.004	1.051**	1.026**
	(0.004)	(0.005)	(0.010)	(0.017)	(0.004)
Cold	0.860**	0.826**	0.930	0.903’	0.943**
	(0.018)	(0.022)	(0.046)	(0.053)	(0.017)
Constant	45,442.6**	33,894.1**	22,839.2**	2,975.3**	2,086.7**
	(818.166)	(707.771)	(1,054.363)	(201.647	(40.978)
Adj. R^2	0.273	0.299	0.416	0.471	0.263
Clusters	3,656	3,653	3,525	2,186	3,655
N	161,208	161,139	155,765	97,342	161,201

(b) Intensive margin

	Total	Car	Public Transport	Bicycle	Walking
Pricing	0.986	0.973’	1.025	0.989	1.040*
	(0.014)	(0.016)	(0.037)	(0.052)	(0.017)
Information	1.009	0.986	1.093*	1.001	1.014
	(0.014)	(0.015)	(0.040)	(0.048)	(0.017)
Difference	0.977	0.987	0.938’	0.988	1.025
	(0.014)	(0.016)	(0.032)	(0.056)	(0.017)
Precipitation	0.999	1.000	0.997	0.999	0.995**
	(0.001)	(0.001)	(0.002)	(0.004)	(0.001)
Heat	1.047**	1.064**	1.023*	1.026**	1.029**
	(0.004)	(0.004)	(0.009)	(0.009)	(0.004)
Cold	0.857**	0.845**	0.943	0.945	0.941**
	(0.017)	(0.020)	(0.041)	(0.042)	(0.016)
Adj. R^2	0.280	0.298	0.486	0.647	0.261
Clusters	3,656	3,651	3,373	1,541	3,654
N	154,520	125,452	55,889	12,645	135,651

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.12: ATE on mode distance share and congestion per km

(a) Mode Distance Share

	Car	Public transport	Bicycle	Walking
Pricing	0.969**	1.047*	1.178*	1.048’
	(0.008)	(0.024)	(0.080)	(0.027)
Information	0.983*	1.045*	1.067	1.017
	(0.007)	(0.022)	(0.067)	(0.025)
Difference	0.986’	1.001	1.104	1.031
	(0.008)	(0.021)	(0.076)	(0.026)
Precipitation	1.002**	0.995**	0.980**	0.996*
	(0.000)	(0.001)	(0.004)	(0.001)
Heat	1.005*	0.969**	0.993	1.015*
	(0.002)	(0.005)	(0.013)	(0.006)
Cold	0.969**	1.033	0.979	1.019
	(0.010)	(0.023)	(0.042)	(0.021)
Adj. R^2	0.056	0.201	0.303	0.121
Clusters	3,653	3,525	2,186	3,655
N	154,483	149,507	93,927	154,518

(b) Congestion per km

	Congestion (car)	Congestion (PT)	Speed (car)	Speed (PT)
Pricing	0.930**	0.977	1.001	1.010
	(0.019)	(0.040)	(0.006)	(0.013)
Information	0.955*	0.950	1.001	1.023’
	(0.020)	(0.040)	(0.006)	(0.013)
Difference	0.974	1.028	1.000	0.986
	(0.020)	(0.041)	(0.006)	(0.012)
Precipitation	0.999	0.997	0.999’	0.999
	(0.001)	(0.002)	(0.000)	(0.001)
Heat	0.950**	0.948**	1.017**	1.007*
	(0.004)	(0.008)	(0.001)	(0.003)
Cold	1.073**	1.077	0.960**	0.996
	(0.028)	(0.059)	(0.008)	(0.015)
Adj. R^2	0.028	0.038	0.189	0.294
Clusters	3,651	2,517	3,651	3,373
N	125,452	48,295	125,453	55,889

Notes: **: $p < 0.01$, *: $p < 0.05$, ’: $p < 0.1$. Standard errors in parentheses and clustered at participant level.

Table A.13: ATE on probability to travel (extensive margin)

	Total travel	Car	Public Transport	Bicycling	Walking
Pricing	1.000	0.970**	1.062*	1.133*	1.006
	(0.005)	(0.010)	(0.027)	(0.070)	(0.008)
Infomation	0.998	0.980*	1.047'	1.011	1.004
	(0.005)	(0.009)	(0.025)	(0.059)	(0.007)
Difference	1.002	0.990	1.014	1.121'	1.001
	(0.005)	(0.010)	(0.026)	(0.072)	(0.008)
N	101,787	151,263	155,620	97,342	151,458

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.14: ATE on departure time for car trips

	Overall	Morning	Evening
Pricing	0.381	-4.680*	1.850
	(2.525)	(1.964)	(2.181)
Info_post	-3.179	-1.745	-0.083
	(2.499)	(1.989)	(2.207)
Difference	3.560	-2.934	1.933
	(2.528)	(2.096)	(2.242)
Adj. R^2	0.054	0.204	0.116
Clusters	2,886	2,885	2,886
N	285,077	102,410	182,665

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.15: Trends in distance travelled

	All Travel	All Travel	Car	PT	Bicycle	Walking
Info	0.099	0.242	-0.093	0.291	-0.067	0.111’
	(1.107)	(1.461)	(1.329)	(0.835)	(0.115)	(0.067)
Pricing	2.456*	2.546’	0.694	1.689’	0.034	0.129*
	(1.133)	(1.510)	(1.348)	(0.908)	(0.134)	(0.065)
Study Day	0.042	0.027	-0.046	0.066*	-0.001	0.007*
	(0.038)	(0.053)	(0.046)	(0.032)	(0.004)	(0.002)
Pricing x Study Day		-0.006	0.012	-0.011	-0.003	-0.004
		(0.063)	(0.055)	(0.037)	(0.005)	(0.003)
Info x Study Day		0.051	0.066	-0.013	0.001	-0.004
		(0.063)	(0.055)	(0.037)	(0.004)	(0.003)
Calendar day FE	☑	☑	☑	☑	☑	☑
Adj. R^2	0.005	0.005	0.005	0.007	0.007	0.012
N	83,081	83,081	83,081	83,081	83,081	83,081

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

In this section, we describe the study in more detail.

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.[20] 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 3,400 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

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

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.

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 the 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 opinion 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.

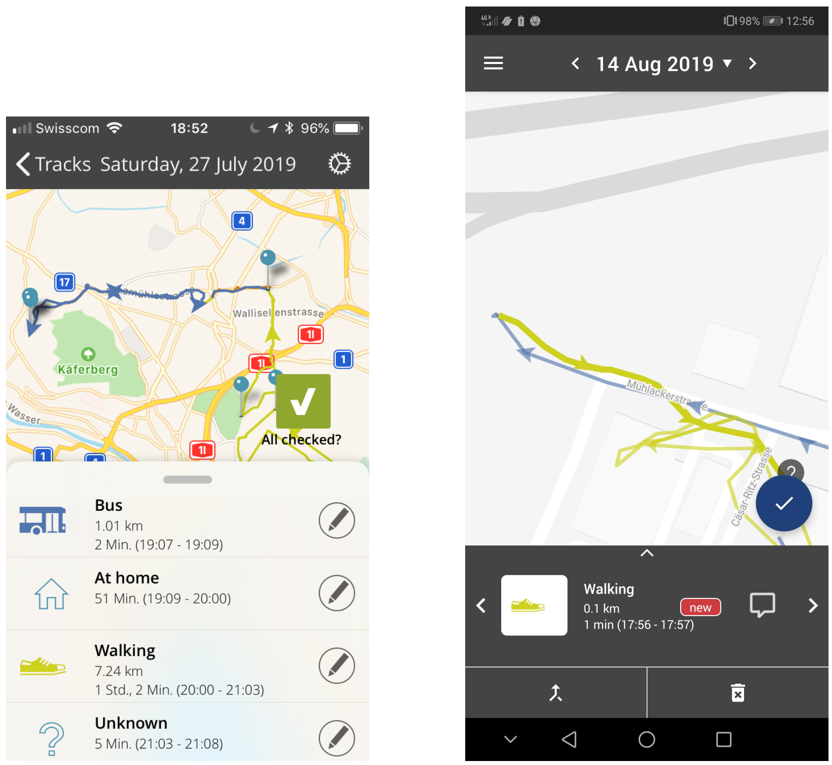
B.4 Tracking

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 can 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.

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



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 in question has made at least one correction during the course of the experiment.[21]

Table B.1 provides the accuracy rates using this assumption. There is 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.

[21] Among the participants that used the validation functionality (85.7%) functionality, 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 will therefore assume that these participants did not use or understand the validation interface correctly and and these participants are 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

	% Correct			
Mode	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

	Confirmed mode									
	Airplane	Bicycle	Boat	Bus	Car	Rail	Tram	Walk	Other	**Total**
Predicted										
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**

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 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.

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

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 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.

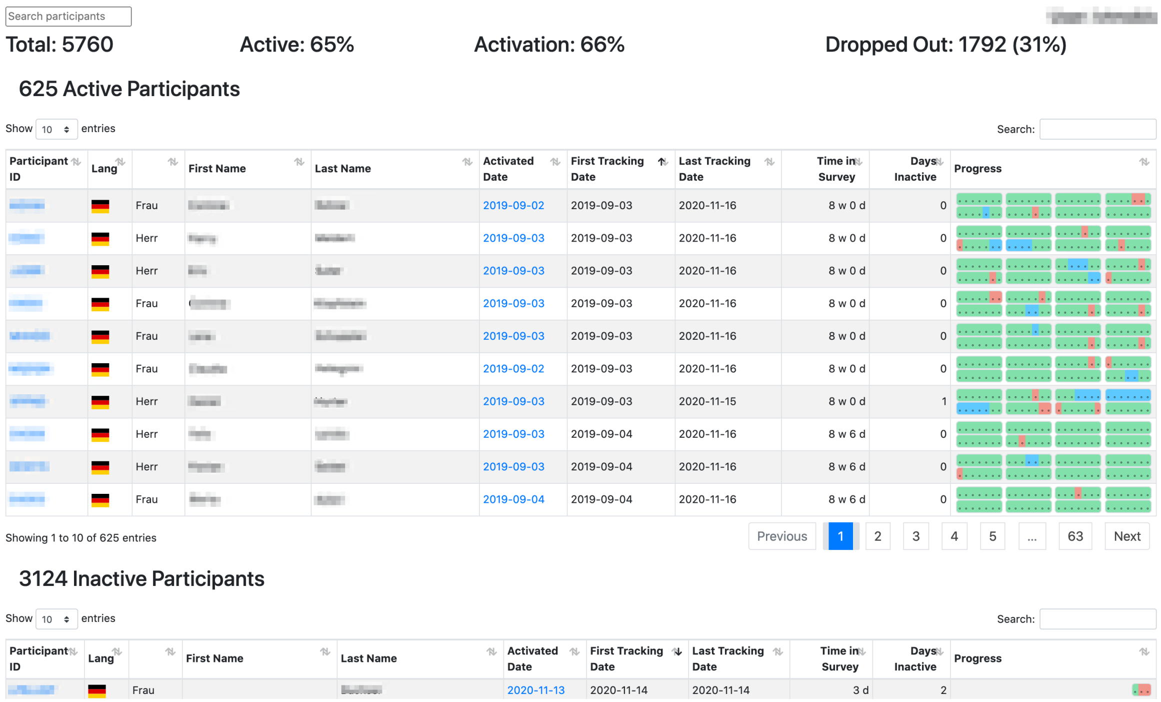
A project website was created to support people invited to the MOBIS study. The website contained links to the introduction survey and the tracking study registration, a project description, information for study participants (including a general information sheet, instructions for the tracking app, 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

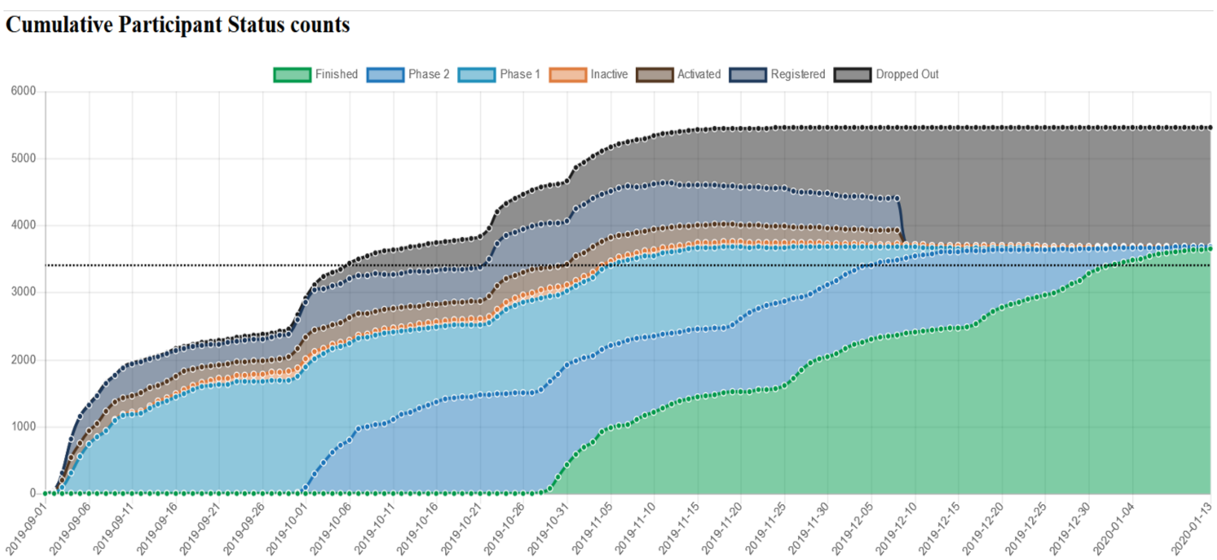
To explore the retention rate of participants in the tracking phase, we performed a survival analysis on the duration of tracking in the study. First, a Kaplan-Meier approach (see Figure B.4) shows the impact of the treatment on the length of time which participants would track. Participants who were automatically dropped out after phase 1 due to poor tracking compliance but were still tracking at the end of phase 1 were censored (marked

Figure B.2: Overview page of participants



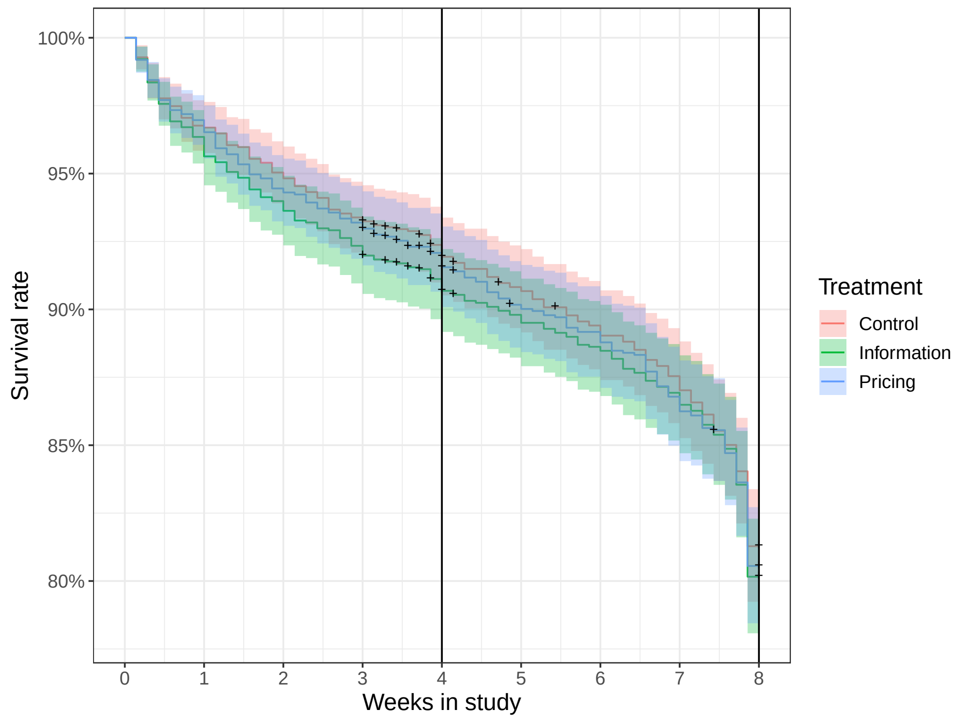
Note: This screenshot was taken after the conclusion of the study, and the participants counts do not reflect the real status during the study.

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



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.4: Kaplan-Meier survival curve by treatment group



Note: 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

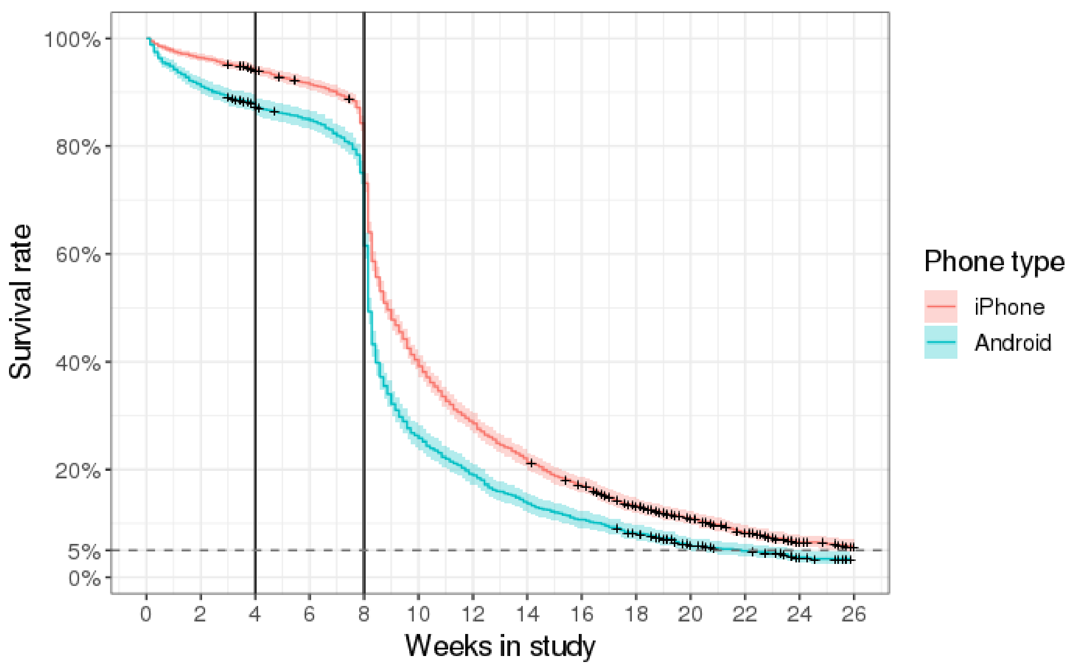


Table B.4: Cox porportional-hazard model

	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.