University of Basel
Faculty of
Business and Economics

WWZ

April 2024 (replaces version of October 2021 and May 2022)

Pigovian Transport Pricing in Practice

WWZ Working Paper 2021/11

Beat Hintermann; Beaumont Schoeman; Joseph Molloy; Thomas Götschi; 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†

April 19, 2024

Abstract

We implement Pigovian transport pricing in a field experiment in urban agglomerations of Switzerland over the course of 8 weeks. The pricing considers external costs from climate damages, health outcomes and congestion and varies across time, space and mode of transport. The treatment reduces the external costs of transport of the treated individuals by 4.5% in the short run. The main underlying mechanism is a shift away from driving towards other modes, such as public transport, walking and cycling. Providing information about external costs alone changes behavior of altruists, but not for the whole sample. We estimate the welfare improvements from such a policy to be around 140 US dollars per person and year, which is twice as large as the effects of a fuel tax that generates the same revenue.

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

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

*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 high costs due to competition with other land use. Furthermore, greater road capacity induces demand and does not alleviate congestion (Duranton and Turner, 2011). The transport sector is also among the largest contributors of local air pollution (EEA, 2019) and greenhouse gas emissions (Creutzig et al., 2015), which have plateaued as gains in efficiency have been neutralized by increases in distance traveled (IEA, 2020).

Congestion, climate damages and health effects constitute the most important external costs of transport.[1] Whereas the private costs of transport, such as the purchase of fuel or a transport pass, have been shown to influence individual transport choices , the external costs are borne by society at large and are typically not reflected in the decision about where, when and how to travel. 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. The main modes considered are car, public transport, cycling and walking.

Our study employs a randomized controlled trial (RCT) design embedded within a tracking study, which allows for unbiased estimates of treatment effects as well as an analysis of the underlying mechanisms. Our sample consists of people living in urban agglomerations of the German- and French-speaking regions of Switzerland. The pricing affects all modes and is implemented by providing the participants with a personalized budget, from which the external costs of their transport choices are subtracted during a period of four weeks. The average short-term effect of the treatment is a reduction in the external costs of transport of 4.5%. The reduction in the external costs is due to a mode shift away from driving and, to a lesser extent, due to a shift in departure times. Car owners, people living in rural areas, and those below 30 years of age respond more strongly than the average. The response is driven by those participants that understood the concept of external costs of transport.

To differentiate the pricing effect from a pure information effect, the experiment includes a second treatment arm in which the participants are provided with the exact same information about the external costs of transport as the pricing group, but without having to pay anything. The results suggest that providing information *per se* results in a significant

[1] Another category of transportation costs are the infrastructure costs, which are considered to be fixed. For a recent discussion of the definition of the external costs of transport, see CE Delft (2019).

reduction for participants with an above-average score for altruism, but not for the sample as a whole. The differential effect between the pricing and information groups can be interpreted as the causal effect of adding a financial incentive for people who are informed about their external costs of transport. This "pure pricing" effect is particularly important in the reduction of congestion externalities. Our results thus imply that information and monetary incentives each play an important and separate role in explaining the total effect of transport pricing. We examine this more deeply by engaging in a mediation analysis in which a participant's driving distance serves as the mediator. We find evidence for a significant effect of the information treatment via this mechanism.

Last, we estimate the welfare effects from Pigovian transport pricing by constructing non-chosen alternatives and estimating a discrete-choice model on the trip level. We use the resulting preference parameters to estimate the monetized utility loss from our pricing. The total welfare effect is the sum of this utility loss, the raised revenue and the reduction in external costs. In our central estimate, we compute a (short-run) welfare gain of CHF 140 per person and year. Due to additional substitution possibilities, the long-run effect is likely larger. Based on our model, we find that second-price policies such as a fuel tax or a perimeter pricing scheme capture at most 50% and 87% of this effect, respectively.

In most real-world settings, the external costs of transport have been 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). From an economics point of view, 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 theoretical foundations for efficient transport pricing were laid by Pigou (1920), Knight (1924) and Vickrey (1963). In first-best pricing, the full marginal external cost is charged to all users, who will then internalize the external costs when making their private transport choices. Beheshtian et al. (2020) propose a multi-modal network management scheme for congested transportation systems based on insights from efficient electricity market mechanisms. In second-best, the pricing mechanism is also guided by the principle of marginal external costs, but the implemented scheme is simplified (Verhoef, 2000; Small et al., 2007). In a recent paper, Almagro et al. (2024) compute the optimal congestion tax, transit price and frequency for a budget-constrained city and calibrate it using data from Chicago.

The most prevalent examples of price-based instruments in the transport sector are fuel taxes, road tolls and registration fees. They are usually imposed to recover only the cost of road construction and maintenance, and thus typically do not reflect the full external costs of transport (Parry and Small, 2005; Parry et al., 2007). Congestion charges can act as an effective way to internalize some of the congestion costs of driving (Small, 2008), and several cities have introduced fees for driving into the city center at certain times. However, since these fees tend to be fixed, they cannot fully address the time-varying nature of congestion. Furthermore, congestion charges usually target only one transport mode and ignore other external costs, which raises concerns about efficiency and equity within the transport sector regulation. In this study, we implement first-best pricing that includes all relevant modes and thus sidestep the various issues that arise when departing from the Pigovian approach.

Besides these conceptual considerations of how transport should be priced, there is also a growing literature about the empirical effects of such pricing. Previous research includes estimates of the aggregate effect of 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). Evidence from the congestion charges in Norway and Milan suggests that they were effective in reducing not only congestion but also local air pollution (Isaksen and Johansen, 2021; Gibson and Carnovale, 2015).

In the absence of such real-world interventions, a number of experimental studies has exposed study participant to an artificial pricing scheme (for a review, see Dixit et al., 2017). For example, studies in Denmark and Australia exposed drivers to different peak and off-peak charges (Nielsen, 2004; Martin and Thornton, 2017), and commuters using public transit in Singapore were exposed to rewards and social comparisons with the aim of shifting demand towards off-peak times (Pluntke and Prabhakar, 2013). Most of these studies focused on a single mode of transport and could therefore not measure modal shifts. A notable exception is the "Spitsmijden" experiment in the Netherlands, in which commuters responded to financial rewards by shifting departure times, switching to other modes of transport and working from home (Ben-Elia and Ettema, 2011).

While highly informative in their respective contexts, all of these studies used a before-vs-after design. The identification strategy is then based on the assumption that no other important determinants of transport changed as the prices were introduced. By including a control group that is never treated, we can control for time-varying shocks and thus causally identify the effect of our treatments.

To the best of our knowledge, only four previous RCTs have been published to identify the impact of financial incentives on transport choices. The first is by Rosenfield et al. (2020), who carry out an experiment involving 2,000 employees at the Massachussetts Institute of Technology. They find no statistically significant effects of a parking fee on parking events. Goldszmidt et al. (2020) use exogenous variation in the price and waiting time for Lyft customers in the US to identify the value of time, and that this depends on market factors such as the proximity to a transit stop. Christensen and Osman (2023) carry out an experiment with Uber clients in Egypt and report that a 50 % discount quadruples demand, some of which comes from a substitution away from public transport (especially for women). However, none of these experiments was able to directly monitor travel for non-car modes, and the evidence for modal shift is thus inferred from the reduced demand for driving only. Finally, Kreindler (2023) uses an experiment to measure the effect of a departure time charge and a zonal price on drivers in Bangalore and computes significant treatment effects using a smartphone app similar to ours. To the best of our knowledge, MOBIS is the first explicitly multi-modal RCT of a pricing intervention in the transport context.

Behavioral change could also be achieved by means of non-financial interventions, which may be easier to implement than prices or taxes. A number of studies have investigated the effect of non-financial interventions in the transport sector (see Möser and Bamberg, 2008, for a review), and some recent papers have used tracking apps to test the effect of informational interventions, however, these are based on small samples (Maerivoet et al., 2012; Carreras et al., 2012; Bothos et al., 2014; Jariyasunant et al., 2015). Kristal and Whillans (2020) use a large-scale RCT to examine the effect of information-based measures on car pooling but find no effect.

Our paper makes several contributions to the literature. First, we show that it is possible to compute person-, time- and location-specific taxes and apply them in the field (proof of concept). Second, by implementing this pricing scheme in an RCT involving a representative sample of the population living in large urban agglomerations, we obtain credible information about the short-run behavioral response to transport pricing, including modal substitution. Third, by estimating a structural model of transport demand, we can compute the welfare effects of our first-best transport policy, as well as second-best versions such as fuel or congestion taxes. Last, as we apply an information-only treatment within the same experiment, our study 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, the data and the computation of the external costs of transport. Section 5 contains the reduced-form results and section 6 the results from our structural model. Section 7 concludes.

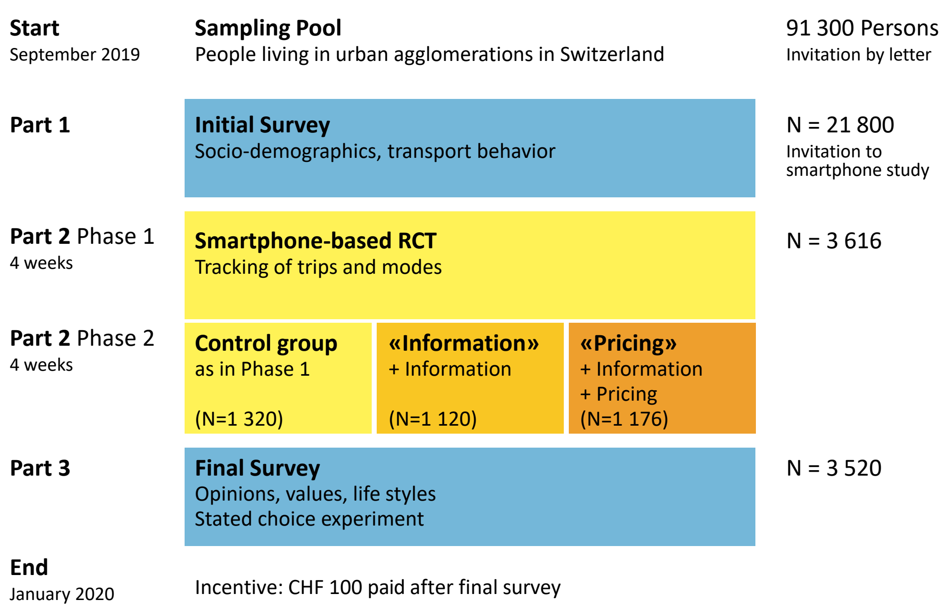
2 The MOBIS experiment

In the following, we describe the main design elements of the Mobility in Switzerland (MOBIS) project. For more detailed information about the study design (including recruitment and attrition rates), we refer the interested reader to Molloy et al. (2023) and Appendix B.

2.1 Study design and sampling

The sample for MOBIS project was recruited among individuals living in the main urban agglomerations in the German- and French-speaking regions of Switzerland. Figure 1 provides an overview of the study design. We contacted a representative sample of 91,300 people by letter and invited them to participate in the study. The letters were written in German and French (depending on the region), with an English translation on the back page. The majority of the addresses were randomly selected and provided by the Federal Office of Statistics, which maintains a comprehensive registry of inhabitants; the remainder was obtained from a private vendor.[2]

Figure 1: Design of the MOBIS experiment



The study contained three parts. The first consisted of an initial online survey, which was completed by close to 22,000 respondents. It contained questions about travel behavior and socio-demographics and served as a screening mechanism. To be invited for the second part (the RCT), respondents had to use a car on at least 2 days per week but could not be professional drivers. Around 11,000 respondents from the initial survey qualified for the RCT, and a total of 5,466 registered for it. However, not everyone who registered actually started tracking. The third part consisted in a second survey.

Participation in the RCT required the tracking of daily travel by means of a smartphone app over a period of 8 weeks. All participants of the RCT were offered an incentive payment

[2]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 desired number of participants, we purchased an additional 31,000 addresses from a private marketing firm.

of CHF 100, which they received after completing the tracking and a final survey.[3] The recruitment took place on a rolling basis between August and November 2019. Once the participants registered their first track on the app, they became part of the RCT sample.[4] The participants knew that they were being recruited for a transport-related tracking study, but we were careful not to mention an experiment nor external costs. Once the study was concluded, all participants were informed about having taken part in a research experiment.[5]

After 4 weeks, the participants in the RCT were randomly assigned to either the control or one of two treatment groups, with a probability of one-third each. Table 1 shows that randomization worked well, with most variable being balanced across the three groups in the RCT sample. However, because of the "double" self-selection (first into the survey and then into the tracking part of the study) and the driving requirement, a careful look at the composition of this sample relative to the general population is warranted.

The table shows summary statistics of some key socio-demographic variables for the sample that filled in the introduction survey, the RCT sample, and the Mobility and Transport Microcensus (MTMC), which is a representative travel diary survey of the Swiss population undertaken by the Federal Office of Statistics and the Federal Office of Spatial Development (2017). To provide a meaningful comparison, we restrict the MTMC sample to the same age range (18-65) and geographic area as our study. We see that the respondents of the introduction survey are similar to the MTMC population. The largest differences are in terms of the share of young adults aged 18-25, education and nationality.[6] Employment, gender, household size, income, language and degree of urbanisation are similar.

The tracking sample has a slightly higher employment rate, more students, and fewer one-person households than the general population, but is similar along most other socio-demographic characteristics. The share of "suburban" residents somewhat larger, which is most likely due to the car driving requirement for participation in the study.[7] Furthermore, the percentage of people that do not have access to a car is lower in the RCT sample than in the MTCM, most likely because we conditioned participation on regularly driving. For

[3]At the start of the study (September 2019), one Swiss Franc (CHF) corresponded to 0.92 Euro and 1.01 US Dollars.

[4]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; see Molloy et al. (2020, 2021); Hintermann et al. (2023).

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

[6]We believe this due to the fact that our recruitment was based on letters and online surveys, whereas the MTMC is based on targeted telephone interviews that include translators when necessary. In contrast, people who are not fluent in German, French or English likely disregarded our invitation.

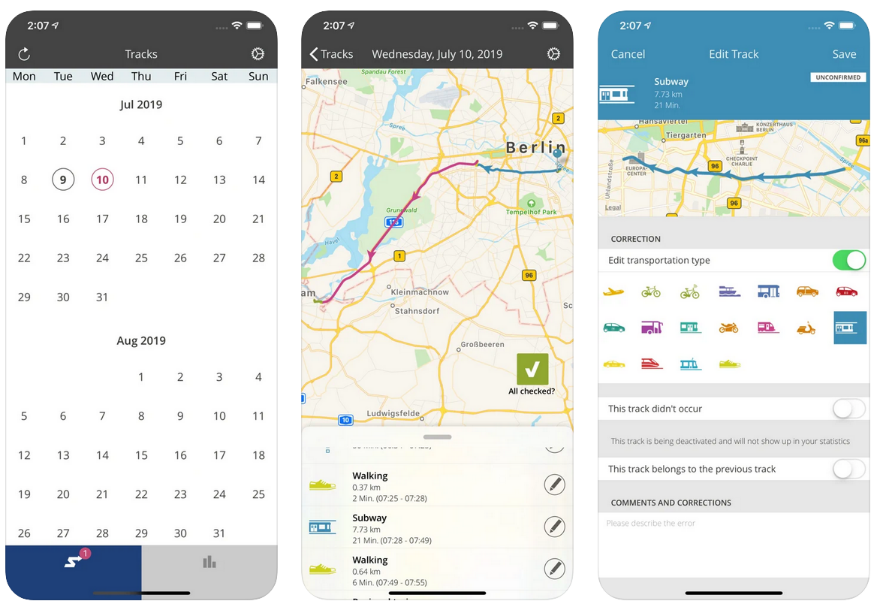
[7]The "urbanization" variable is constructed by allocating participants' home postcodes one of three degrees of population density: urban, suburban, and rural. These definitions are based on the Swiss Federal Statistical Office's definitions, which is partly based on the accessibility of road and public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

the same reason, the share of people who have a full PT subscription is lower. We discuss the implications of our sample selection procedure for external validity in section 6.

2.2 Tracking app

The participants in the tracking-part of the study agreed to download the tracking app "Catch-My-Day" on their smartphones. Catch-My-Day is a location tracker for iOS and Android, which uses the location services of the respective operating system. The GPS tracks are stored on the phone and uploaded to the Motiontag analytics platform, where trip stages are identified and travel modes and activities are imputed based on a machine learning algorithm. For each stage, the associated external costs of transport were computed based on cost factors published by the Swiss Government (see section 2.3). Participants were able to review and correct the mode assignment.

Figure 2: The Catch-my-Day interface



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

Figure 2 shows three interfaces of Catch-my-Day. The app provides a best guess of the travel mode for each stage. The participants could then confirm the imputed mode or correct it. This confirm-correct procedure was optional but participants were informed that this would be appreciated.[8] Around 79% of the stages were confirmed by the participants,

[8] 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 et al., 2016; Nikolic and Bierlaire,

Table 1: Demographic information for the MOBIS sample

Variable	Level	Intro	Tracking					MTMC
			Control	Info	Pricing	Info. diff.	Pricing diff.	
Age	[18, 25]	20.1	18.0	19.6	19.4	-0.802	-0.579	14.3
	(25, 35]	19.4	17.7	18.4	16.6			21.4
	(35, 45]	19.9	22.0	20.8	24.1			22.6
	(45, 55]	21.6	22.8	23.9	22.4			23.7
	(55, 65]	19.0	18.4	16.2	16.1			17.9
Education	Mandatory	9.2	8.0	5.2	6.8	0.009	0.004	13.8
	Secondary	43.3	47.4	49.3	48.2			47.5
	Higher	47.5	44.6	45.6	45.0			38.7
Employment	Employed	68.7	72.7	71.4	70.0	0.050	0.108’	68.8
	Self-employed	7.3	6.2	5.4	7.1			8.8
	Apprentice	1.9	1.9	1.6	1.6			2.2
	Unemployed	4.4	3.3	3.9	4.6			3.9
	Student	9.3	7.5	8.6	7.7			3.0
	Retired	2.5	3.5	2.7	3.6			3.6
	Other	5.9	4.8	6.5	5.4			9.7
Gender	Male	48.9	50.5	50.2	49.4	-0.003	-0.011	49.4
	Female	51.1	49.5	49.8	50.6			50.6
Household size	1	15.5	11.2	11.5	12.1	-0.035	0.021	18.3
	2	31.7	30.2	31.4	29.0			32.0
	3	20.5	23.0	21.3	19.9			19.9
	4	23.6	25.2	27.7	29.7			20.7
	5 or more	8.6	10.3	8.1	9.4			9.1
Monthly household inc.	4 000 CHF or less	12.2	6.6	8.3	7.1	0.021	0.091	8.8
	4 001 - 8 000 CHF	29.4	31.1	29.6	27.2			31.4
	8 001 - 12 000 CHF	24.5	28.0	30.0	30.2			24.6
	12 001 - 16 000 CHF	12.1	15.5	13.9	14.5			11.7
	More than 16 000 CHF	8.0	9.8	9.4	10.6			8.4
	Prefer not to say	13.8	9.0	8.8	10.3			5.8
	Don’t know							9.2
Language	German	62.7	66.6	65.3	66.5	0.017	0.004	69.5
	French	28.6	25.9	26.4	26.3			26.5
	Italian							4.0
	English	8.7	7.5	8.3	7.2			
Nationality	Switzerland	78.1	81.6	80.7	82.6	-0.008	0.010	69.5
	Other	21.9	18.4	19.3	17.4			30.5
Residential setting	Urban	75.0	64.0	64.5	64.8	-0.001	-0.014	77.4
	Suburban	18.1	27.9	27.0	27.7			16.6
	Rural	6.8	8.1	8.5	7.5			6.0
Access to car	Yes	61.0	87.2	88.0	88.0	0.005	0.010	69.7
	Sometimes	15.5	11.6	10.5	11.1			22.7
	No	23.5	1.1	1.5	0.9			7.5
Full PT subscription	Yes	37.2	21.9	25.0	25.3	0.031’	0.034’	34.5
Half fare PT subscr.	Yes	47.6	49.3	49.3	48.4	0.000	-0.009	37.6
No PT subscription	Yes	26	33.7	32.5	33.7	-0.012	0.000	37.9
Access to bicycle	Yes	68.5	72.9	72.2	69.5	-0.002	-0.074*	70.1
	Sometimes	4.1	4.4	5.5	3.8			8.8
	No	27.4	22.7	22.3	26.7			21.1
N		20,783	1,174	1,187	1,134			21,399

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

and 5.1% of the modes were corrected. The database stores both their correction and the original algorithmic imputation. The possibility for mode correction increases the accuracy of the mode detection, but it also introduces a possibility for "gaming" the experiment. As discussed in section 6, we believe that this did not happen to a significant degree.

The following modes are detected the by Catch-my-Day app: Airplane, 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. For the analysis, we retained only trips within Switzerland coded as car (car, car sharing and taxi/Uber), public transport (intercity, regional and local trains, bus, tram), cycling and walking. All other modes were excluded.

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. This algorithm achieved an overall accuracy of over 90% (see Molloy et al., 2023, Table 8).

2.3 The external costs of transport

The assessment of the external costs of transport is based on the cost computations carried out by the Swiss Federal Office of Spatial Development (ARE). The external costs are defined as those that are not yet internalized in the existing framework of transport-related taxes on fuel, vehicle registration fees etc. The external costs of transport can be grouped into climate-related emissions, congestion, and health. The latter category includes health effects associated with local pollution and noise, accident-related health costs and also physical benefits from active transport. Pricing these external costs into individual trips will lead to the full internalization of external costs of transport in a Pigovian sense.[9] Throughout the paper, we focus on *marginal* external costs, i.e., the costs associated with an individual trip conditional on the vehicle having been purchased and the infrastructure having been built. We therefore abstract from the life-cycle emissions associated with producing cars and building road and rail infrastructure, which, at the time an individual chooses to make a trip, have been generated already.

The health, emissions, noise and congestion costs of the mobility behavior were computed

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.

[9] The costs of road maintenance is covered by existing fuel taxes and are therefore already internalized. According to ARE's calculations, private motorized transport generated a total cost of CHF 52,486 million in 2019. Of this, CHF 44,969 million were paid for by road users (e.g., in the form of fuel and taxes), whereas the remaining CHF 7,517 million were borne by society at large (Federal Statistical Office, 2022, Fig. G4). These uncovered costs are the sum of health costs (due to local air pollution, noise and accident-related costs), congestion and climate change.

on the recorded daily trips using an automated data pipeline. Data collected from the online introduction survey (e.g., engine type and size) was incorporated into the data processing pipeline to improve the imputation.

For the calculation of external costs associated with driving, a partial-equilibrium approach described in Molloy et al. (2021) was used. Briefly, the recorded GPS tracks were aligned to the road network using Graphhopper (Karich and Schröder, 2014) and processed using modules developed on top of the MATSim framework to calculate the external costs of congestion and emissions. The emissions factors were taken from the HBEFA database (version 3.3), and applied using the MATSim emissions module (Hülsmann et al., 2011; Kickhöfer et al., 2013). For congestion, an average marginal cost approach incorporating spillback effects and flow congestion was applied, based on the work of Kaddoura (2015).[10] 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 a social cost of carbon of CHF 136 per tCO_2, a value of travel time of CHF 25.77 per hour, CHF 515 (1,358) per kg of PM_10 in rural (urban) areas and 7,109 CHF per ton of NO_x (see Molloy et al., 2021, Table 2). The costs associated with driving vary over time and space mainly due to changing levels of congestion, but also due to different emission factors depending on speed and different exposure of the population to local air pollution (urban vs. rural).

For modes other than driving, the per-km values presented in Table 2 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 reduced morbidity as a consequence of physical activity (Götschi et al., 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 net external costs in the experiment.[11]

The marginal external cost of public transport per person-km decreases as the occupancy rate increases. On the other hand, crowding affects the willingness to pay for public transport and can be seen as a form of congestion in public transport, and delay in some circumstances (Tirachini et al., 2013). Crowding effects are extremely heterogeneous, both spatially and temporally (VBZ, 2017). However, there is insufficient data to compute crowding at a similar level of detail as for congestion. As a practical solution, a peak-hour pricing scheme

[10] An alternative way to proceed is the approach by Yang et al. (2020), who exploit a natural experiment to empirically estimate the causal relationship between traffic density and speed in Beijing. This allows them to compute the optimal congestion charge for that city.

[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 2: Monetary costs per person-km (in CHF) used in the MOBIS experiment

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

Notes: The values for public and active transport are based on NISTRA (Federal Roads Office, 2017). Congestion costs for public transport were only applied for congested links (see text). Negative costs indicate an external benefit. The external costs of driving vary over time and space and were computed within MATSim (Molloy et al., 2021).

was developed for the purpose of the study. The peak windows were set as 7:00 to 9:00 and 17:00 to 19:00 and not adapted for regional variation in working patterns. The peak surcharge of 0.10 CHF/km was applied to transit stages between any two municipalities which experienced a peak hour demand that exceeded offpeak demand (8:00 to 17:00) by more than a factor of 3. The peak hour windows and the affected municipality-pairs were determined using the MATSim scenario output for Switzerland (Bösch et al., 2016). A municipality could also be paired with itself if the above criteria were met and the direction of the peak hour flow was not considered. If the trip was partially in both the peak and off-peak periods, only the proportion of the travel duration that overlapped with the peak period was charged. Throughout the experiment, participants had access to an interactive map which showed them where and when the pricing scheme applied.

In principle, we could have used any pricing scheme and estimated the participants' response to it. We chose the Pigovian rate (i.e., the net marginal external cost pf each trip) for two reasons.[12] First, internalizing the external costs of transport can be motivated on normative grounds. The "information only" treatment could thus be interpreted as providing information on true societal costs about which the participants were likely not perfectly informed. In contrast, introducing a price unrelated to the external costs would be more difficult to justify and therefore would be less likely to lead to behavioral change via an altruistic motive. Second, using the Pigovian rate serves as a policy benchmark. If larger responses are required, the policy maker can choose to exceed this rate (and vice versa), but

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

we believe that knowing the Pigovian rate and people's response to it is useful information.

2.4 Intervention

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. The participants assigned to the control group received these summaries throughout the study.

On tracking day 29, the participants randomly assigned to the "information" and "pricing" groups received an e-mail that informed them about the external costs of transport, how these costs are computed and what the participants could do to reduce them. The e-mail contained a link to a table with average per-km monetized costs by mode (similar to Table 2). To complement this average price information and to provide the participants with an idea about their individual level of external costs, they were also shown a personalized summary of their own external costs from the previous week.[13] For the remainder of the 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 group received the exact 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. These participants were informed that any remaining money in their account at the end of the study was theirs to keep (in addition to the standard incentive of CHF 100 that was paid to everyone). The individualized budgets were computed based on each participants' external costs during the observation period, plus a 20% buffer to allow for the possibility that some participants had to increase their external costs of transport for idiosyncratic reasons.[14] This treatment thus simulated transport pricing based on the monetized marginal external costs of transport.

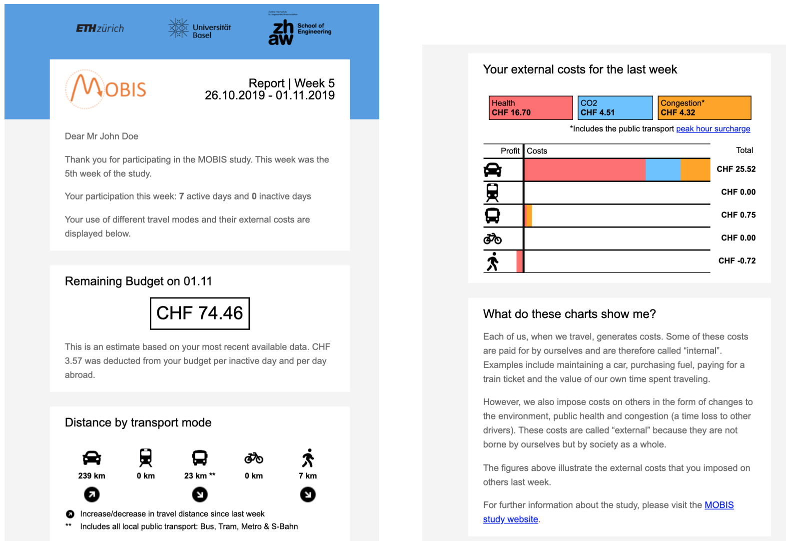
The nested design of the treatments allows for an estimation of the effect of "pure money" in the sense of providing a monetary incentive to individuals that are perfectly informed about their external costs of transport.[15]

[13]To provide participants with ex-ante personalized costs for particular trips was infeasible within the project budget as this would have required a lot of additional programming due to the varying nature of congestion costs. However, the internal part of congestion costs are experienced personally, such that participants likely 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.

[14]We imposed a minimum budget of CHF 50. The average budget was CHF 144, and for some participants it exceeded CHF 700. Participants were told that the budget could not go below zero.

[15]According to economic theory, all prices contain information, such that having an additional "price without information" treatment would not have identified an interesting effect.

Figure 3: Weekly reports by e-mail



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

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 both the Information and pricing groups in the treatment phase, whereas the remaining budget panel (middle module on the left) was presented only to the pricing group. Due to the rolling recruitment into the experiment, participants received these reports on different days of the week.

3 Data and regression framework

3.1 Data preparation

Our starting sample consists of the people for whom we recorded at least 11 tracking days during the observation period; the rest was excluded from the study and never assigned to any group. After receiving the data from the app provider, they were processed using some routine procedures to remove obviously problematic tracking data. Specifically, we remove the data on the person-day-level 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. We remove the first day of tracking (as participants may have started tracking in the middle of the day) as well as day 29 (as it is not clear at what time the participants in the treatment group would open their e-mails). We furthermore restrict the sample to participants that traveled at least 10 km during the observation period. For the main analysis, we remove 71 people who did not record any travel days during the observation period as they do not contribute to the treatment effect; however, these participants are retained for the attrition analysis (see section 6.3). Taken together, these cleaning procedures reduce the original sample of 179,507 person-days from 3,690 participants to 168,362 person-days from 3,616 participants. Of these, 1,120 were assigned to the information group, 1,176 to the pricing group and 1,320 to the control group.

Some participants in the treatment group exhausted their budget before the end of the study and thus no longer had a financial incentive to reduce their external costs. As removing these person-days from the sample could lead to an imbalance between treatment and control groups in terms of mean reversion (see section 6.4), we retain the affected 379 person-days but mark them with a dummy in order not to contaminate the treatment effect.

3.2 Tracking summary statistics

Table 3 shows the summary statistics of the tracking data for all modes combined, including distances, duration, external and private costs. Table A.1 in the Appendix presents the data separately for each mode, and Table A.2 provides a proportional breakdown of the external costs on a aggregation of mode and cost dimension. The external costs are generally dominated by driving. Within driving, the most relevant external costs are associated with health costs; among these, accident costs account for about a quarter, whereas the majority is due to local air pollution and noise. The table shows that most external costs of transport are local, with CO2-related damages amounting to less than 20% of total external costs.

Almost 74% of the recorded distance is traveled by car, and accordingly the vast majority of external costs is associated with driving. Public transport accounts for 21% of overall distance.[16]

3.3 Identification of the treatment effect

Our experiment was designed with the aim of satisfying the four cardinal assumptions to satisfy internal validity for a difference-in-differences design: (i) Statistical independence, (ii) Stable Unit Treatment Value Assumption (SUTVA), (iii) complete compliance and (iv) observability. The random assignment of people to the treatment arms satisfies condition

[16]Public transport is the sum of bus (10.3%), light rail (23.3%), regional train (13.2%), intercity train (48.6%) and tram (4.6%).

Table 3: Tracking summary statistics

Dimension	Outcome	Unit	Observation period Control	Observation period Info	Observation period Pricing	Treatment period Control	Treatment period Info	Treatment period Pricing
External costs	Total	CHF/d	4.50	4.58	4.70	4.24	4.26	4.24
			(5.68)	(5.63)	(5.79)	(5.4)	(5.53)	(5.42)
	Congestion	CHF/d	1.03	1.07	1.14	0.85	0.87	0.90
			(1.59)	(1.58)	(1.69)	(1.43)	(1.52)	(1.51)
	Climate	CHF/d	0.88	0.88	0.90	0.85	0.84	0.83
			(1.29)	(1.28)	(1.29)	(1.23)	(1.28)	(1.23)
	Pollution & noise	CHF/d	1.90	1.94	1.94	1.87	1.88	1.84
			(2.60)	(2.58)	(2.63)	(2.54)	(2.60)	(2.54)
	Health care	CHF/d	0.69	0.70	0.72	0.67	0.68	0.68
			(1.06)	(1.06)	(1.09)	(1.04)	(1.06)	(1.05)
Private costs	Total	CHF/d	22.25	22.65	23.09	22.15	22.14	22.1
			(31.36)	(30.90)	(32.19)	(31.86)	(31.48)	(31.79)
Tracking	Distance	km/d	46.86	47.92	49.34	45.55	47.15	47.25
			(55.32)	(54.63)	(57.28)	(54.18)	(55.81)	(55.10)
	Duration	min/d	92.55	93.31	94.09	88.5	90.49	91.16
			(84.61)	(80.05)	(82.83)	(78.72)	(82.02)	(83.04)
	Tracking days	Nr.	23.64	23.91	23.68	22.75	22.92	22.77
			(3.74)	(3.44)	(3.72)	(5.12)	(4.92)	(5.27)
	Trips	Nr./day	4.71	4.74	4.76	4.53	4.49	4.55
			(3.02)	(3.00)	(2.96)	(2.81)	(2.79)	(2.83)

Notes: Average values per participant over the course of the study (SD in parentheses). "Health care" includes the external costs from accidents from all modes, net of the external health benefits from walking and cycling.

(i) by construction, and since our RCT sample consists of around 3,600 individuals out of an overall population of several million (and the treatment is administered identically across the treated), SUTVA arguably holds too. As we define the treatment as receiving (but not necessarily reading) e-mails from us, compliance was complete too in the sense that no person in the control group received a treatment e-mail, and all people in the treatment groups were sent e-mails. And although we do not observe all people on all days, we show below (Subsection 6.3) that observability does not systematically vary over the treatment assignment nor the determinants of the potential outcomes.

Conditional on these assumptions being met, the average treatment effect (ATE) can be estimated by comparing means between treated and control observations. We aggregate the data to the person-day level and estimate the ATE using the following regression:

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

The dependent variable is the outcome of interest for person $i \in N$ on calendar day $t \in T$ and day of study $s \in (2, ..., 56)$. The main outcome of interest is the total external cost (in CHF per day), but we also run regressions in which the dependent variable is the external cost along a particular dimension (health, climate and congestion), the distance traveled or the average time of departure.

The two difference-in-differences terms DiD^P_{its} and DiD^I_{its} are the products of treatment group and treatment period dummies and are equal to one if the pricing (P) and information treatment (I), respectively, are active for person i on a given day, and zero otherwise. Since we excluded day 29 from the sample (see above), the treatment starts on the 30th tracking day for each participant. Due to the rolling recruitment, this day falls on different calendar days for the participants.

To control for unobserved heterogeneity, we include fixed effects on the person (μ_i), calendar day (μ_t) and day-of-study (μ_s) level. The calendar day FE capture common shocks that affect travel (and thus the associated external costs) for everyone in Switzerland, e.g., due to seasonality, a national holiday or a sports event. The day-of-study FE account for the possibility that respondents change their behavior in response to being tracked (regardless of the group assignment). Importantly, 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. This is similar in spirit to the identification strategy of forming "group-time" averages in Callaway and Sant'Anna (2021), such that treated units are only compared to never-treated units.[17] The error term ϵ_{its} has an expected mean of zero and a variance of σ^2. We allow for a correlation of the error within participants, but not between. Finally, note that the problems involving two-way FE estimators does not apply to our setting as all participants have an equal number of pre- and post-treatment observations.

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 existing information" is their difference, $\alpha^P - \alpha^I$. This value could also be computed by estimating (1) for the pricing group while using the information group as the control. It is therefore a causal ATE in its own right, rather than simply a difference between two coefficients.

Due to the presence of the control group, we do not need to control for any covariates in principle as they are expected to be balanced across groups. However, because our sample is finite and weather information is an important predictor especially for active transport, we enrich our tracking data with temperature (in Celsius) and precipitation (in mm/h) data

[17]Note also that whereas participants started the experiment on different dates, we do not have a "staggered" design in the sense that everyone is treated eventually, and some participants are treated earlier (and thus longer) than others. Regardless of the start date, each person spends exactly 4 weeks in the pre-treatment and another 4 weeks in the treatment phase. The symmetry between pre- and post-treatment observations, per individual, should lead to no one receiving an overall weight in the computation of the ATE.

from MeteoSwiss provided on a 1 x 1 km grid.[18] The weather variables are assigned for each recorded trip based on the weather station nearest to the point of departure. To allow for a nonlinear effect of temperature on travel choices, we define the level of "Heat" and and "Cold" for an observed trip j on day t relative to threshold values:

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

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

The variables t_{jt}^{max} and t_{jt}^{min} refer to the daily maximum and minimum temperature, respectively, recorded in degrees Celsius. We compute the average of the heat, cold and precipitation values across all trips taken by person i on day t and add them as linear control variables to (1). To investigate potential differences of the treatment effect along major socio-economic variables (moderation), we further interact the DiD terms with categorical variables denoting, e.g., gender or income groups (see section 4.2).

For the regressions that use external costs as the dependent variable, we estimate eq. (1) in levels (rather than in logs). This is necessary because the external benefit associated with walking leads to some person-day observations with a negative external cost (i.e., a net benefit). We compute the proportional response by dividing the coefficients (in CHF/d) by the average daily external costs generated by the control group during the treatment period. For regressions in which the dependent variable is non-negative (e.g., distance traveled), we estimate proportional effects directly by using a Poisson Pseudo-Maximum Likelihood (PPML) model. This approach addresses the presence of zeros in the data and the possible presence of heteroskedasticity, which can lead to a bias in log-linearized regressions.[19]

To learn about the mechanisms underlying our effect, we engage in a mediation analysis. As we will see below, an important mediator of the treatment effect on the external costs of transport is the quantity of driving (i.e., car-km per person-day). Following Baron and Kenny (1986), we estimate the following equations:

$$M_{it} = \gamma_0 + \gamma_T T_{it} + \gamma_X X_{it} + \epsilon_{it}^* \tag{4}$$

$$Y_{it} = \beta_0 + \beta_T T_{it} + \beta_M M_{it} + \beta_X X_{it} + \epsilon_{it} \tag{5}$$

Here, M_{it} refers to the mediator, Y_{it} is the outcome variable of interest, T_{it} is the treatment indicator and X_{it} is a vector of controls (in our case, this includes the person, calendar day

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

[19] For a discussion of the advantages of using a Poisson model in the presence of zeros and heteroskedasticity, see Santos Silva and Tenreyro (2006). For estimation, we use an estimator developed by Correia et al. (2019) and Correia et al. (2020).

and study day FE and the weather controls). The identifying assumptions are that ϵ^*_{it} is independent of the treatment status, and ϵ_{it} is independent of both T_{it} and M_{it} (this is also known as the "sequential ignorability assumption"). Given (4)-(5), the Average Direct Effect (ADE) and the Average Indirect Effect (AIE) can then be computed as follows:[20]

$$\text{ADE} = \beta_T \tag{6}$$

$$\text{AIE} = \gamma_T \beta_M \tag{7}$$

The AIE captures the causal effect of the treatment on the outcome variable via the mediator, whereas the ADE measures the "direct" effect (which is the sum of any truly direct effect and the effect of all other mediators).

4 Results

4.1 Average treatment effects

Table 4 shows the ATE on the external costs of travel in CHF per day. The first two columns report the results for the total external costs of transport, with and without controlling for the weather, whereas the next three pairs of columns contain the ATE on the external health, climate and congestion costs. About half of the reduction in external costs is due to a decrease in health costs, followed in magnitude by congestion and then climate costs. Although the weather variables are jointly highly significant, they have no affect on the ATE. For the remainder of the paper, we will therefore refrain from showing both versions.

Figure 4 displays the ATE in proportional terms. There is a statistically significant reduction for all dimensions of external costs, but the effect is particularly large for congestion costs. As we will see below, this likely reflects the fact that congestion can be reduced not only by a mode shift, but also a shift in departure time.

The effect of providing information alone has a negative point estimate in Table 4, but it is not statistically significant for the sample as a whole. The effect of adding a price to information (="Difference) is statistically significant only for congestion. We can think of two possible explanations for this result: People may be more informed about congestion costs than about the other external costs of transport, such that the information is less novel; and second, the altruistic gain from reducing the time that other road users spend driving may be less than the equivalent gain associated with a reduction in climate and

[20] As there is no reason to assume that the outcome variable affects the mediator, an interaction term as proposed by Kraemer et al. (2008) is not needed. Furthermore, as both the mediator and the outcome variable are continuous, no extension in the spirit of Imai et al. (2010) is necessary.

Table 4: Average treatment effects on external costs

	Total ext. costs		Health costs		Climate costs		Congestion costs	
Pricing	-0.225**	-0.226**	-0.116**	-0.118**	-0.038*	-0.038*	-0.071**	-0.071**
	(0.070)	(0.070)	(0.043)	(0.043)	(0.016)	(0.016)	(0.022)	(0.022)
Information	-0.091	-0.096	-0.049	-0.052	-0.020	-0.021	-0.022	-0.022
	(0.067)	(0.067)	(0.042)	(0.042)	(0.015)	(0.015)	(0.021)	(0.021)
Difference	-0.133'	-0.131'	-0.067	-0.066	-0.017	-0.017	-0.049*	-0.048*
	(0.070)	(0.070)	(0.043)	(0.043)	(0.016)	(0.016)	(0.021)	(0.021)
Precipitation		0.001		-0.000		-0.000		0.002
		(0.004)		(0.003)		(0.001)		(0.001)
Heat		0.187**		0.157**		0.057**		-0.027**
		(0.018)		(0.012)		(0.004)		(0.004)
Cold		-0.502**		-0.358**		-0.129**		-0.014
		(0.073)		(0.048)		(0.017)		(0.019)
Adj. R^2	0.233	0.234	0.225	0.227	0.222	0.225	0.265	0.266
Clusters	3,616	3,616	3,616	3,616	3,616	3,616	3,616	3,616
N	168,362	168,362	168,362	168,362	168,362	168,362	168,362	168,362

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. 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. All regressions include fixed effects on the person, calendar-day and day-of-study level.

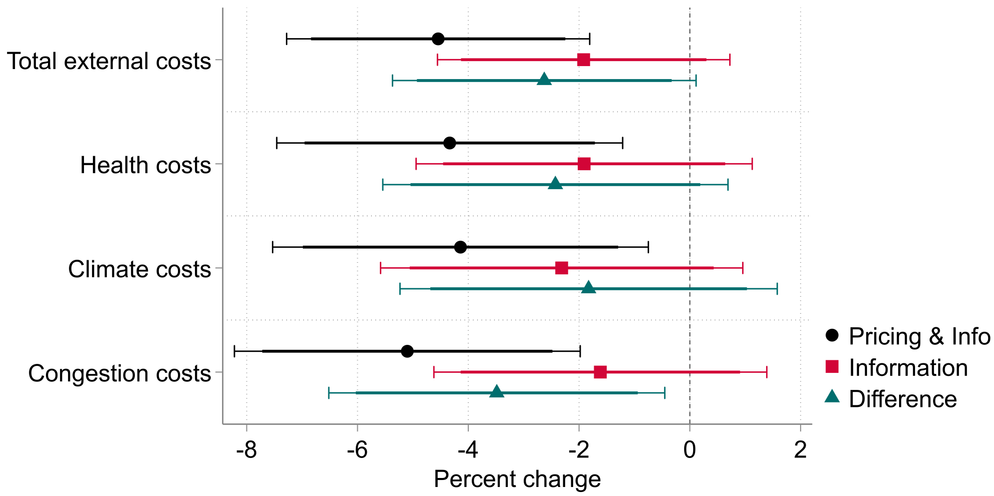
health externalities. In both cases, the price dimension of the intervention carries relatively more weight than the information dimension.

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

To interpret the magnitude of the ATE, we compare the proportional response in external costs to the change in the total transport price, including both private and external costs. For public transport, we use the ticket price as a reference for the private costs.[21] For driving, we directly elicited the private cost from respondents in terms of cents per kilometer. For those that did not answer this question in the final survey, we computed their expected private costs based on information about car size, age and fuel type. This led to an average

[21]For participants that hold a flat-rate public transport pass, we approximated the average cost by applying a discount to the half-fare ticket price. The level of the discount is determined by comparing the cost of a regional PT subscription with the corresponding cost if one were to buy a daily pass on 22 days per month. This procedure returns the average per-km price for public transport, not the marginal cost faced by a participant. However, if transport pricing were introduced in reality, this would be the relevant price increase. The savings implicit in the subscription ranges from 24% in Geneva to 76% in Basel.

Figure 4: Treatment effect on the external costs of transport



Notes: The figure shows the proportional treatment effects for overall travel. They are computed by scaling the regression coefficients in Table 4 by the external cost of the pricing group during the observation period scaled by the temporal change observed for the control group. The bars show 90% and 95% confidence intervals.

cost of 59 cents/km, with an interquartile range of 50-70 cents/km.[22] We abstract from the purchase or rental price of bicycles and set the private cost of cycling and walking to zero.

Given our modeling of private costs, the we obtain a tax-related price increase of 19.3%. The proportional reduction in the external costs is 4.5%. Dividing the latter by the former results in an elasticity of -0.24 that describes the short-term elasticity of external costs in response to a one-percent increase in the costs of transport. Note, however, that this estimate does not identify a fundamental behavioral parameter as it depends on how the relative price change comes about.[23]

Table 5 shows the sensitivity of the results with respect to the inclusion of fixed effects and the presence of a control group. The base model is in column (1). Removing either the day of study fixed effects (col. 2) or the calendar day fixed effects (col. 3) does not significantly change the results; however, when removing both, the ATE more than doubles (col. 4). Controlling for unobserved characteristics that vary over time is therefore crucial

[22]3,178 out of 3,521 respondents answered this question in the final survey. We excluded values below 1, which were presumably meant as francs/km instead of cents/km as specified in the question. To account for unrealistically low or high values (e.g., reflecting the value of a new car rather than the cost per km) we removed the top and bottom 5%. This left us with 2,459 values. Finally, we imputed the missing values based on a linear model associating private costs with information about the age, size and fuel type.

[23]For example, if the same behavioral response were to be achieved by subsidies rather than taxes, the resulting elasticity would be positive. More generally, an infinity of elasticities could be associated with the observed reduction in external costs by 4.5%. The estimate of -0.24 is the special case in which the price change is due to applying the Pigovian rate to all externalities.

for identification. The ATE is significantly over-estimated in the before-vs.-after setting (columns 5-6), because the treatment also absorbs a part of the seasonal effects.

Table 5: Sensitivity to inclusion of fixed effects

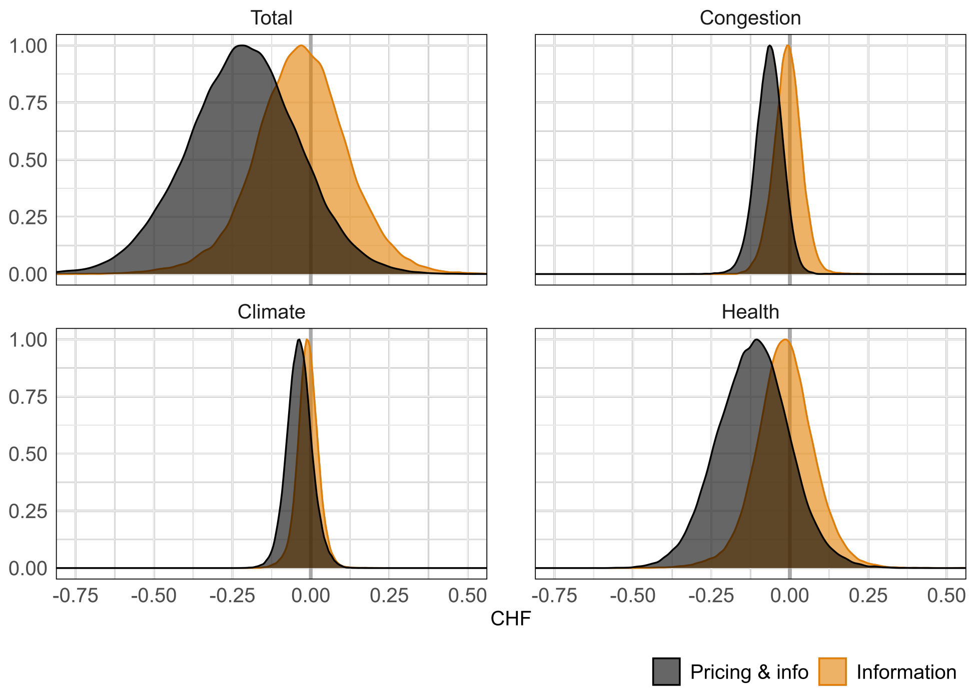
	(1)	(2)	(3)	(4)	(5)	(6)
Pricing	-0.226**	-0.232**	-0.221**	-0.534**		
	(0.070)	(0.062)	(0.070)	(0.053)		
Information	-0.096	-0.100'	-0.087	-0.402**		
	(0.067)	(0.059)	(0.067)	(0.050)		
Post					-0.299**	-0.585**
					(0.104)	(0.056)
Proportional effect	-0.045**	-0.046**	-0.044**	-0.107**	-0.064**	-0.124**
	(0.014)	(0.013)	(0.014)	(0.010)	(0.021)	(0.011)
Elasticity	-0.236**	-0.241**	-0.230**	-0.557**	-0.313**	-0.612**
	(0.074)	(0.065)	(0.074)	(0.054)	(0.105)	(0.055)
	(0.086)	(0.077)	(0.087)	(0.067)	(0.105)	(0.055)
Person FE	☑	☑	☑	☑	☑	☑
Cal. day FE	☑	☑	☐	☐	☑	☐
Study day FE	☑	☐	☑	☐	☐	☐
Adj. R^2	0.234	0.234	0.230	0.229	0.228	0.224
Clusters	3,616	3,616	3,616	3,616	1,176	1,176
N	168,362	168,362	168,362	168,362	54,625	54,626

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 and zero otherwise. In columns (1)-(4), the proportional values and elasticities are derived using the external costs and price increases of the control group. For the before-vs.-after analysis in columns (5)-(6), the daily external cost and price increase was computed for the pricing group during the observation phase. All regressions include weather controls.

4.2 Effect heterogeneity

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

Figure 5: Distribution of conditional effects



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

Figure 5 shows the distribution of the conditional treatment effect, both for the pricing and the information treatments. The relative variable importance derived from the algorithm is shown in Figure A.2. To interpret this measure, we also included a continuous and a discrete random variable. The variables with a higher importance ranking than these random variables can be treated as likely candidates to explain the effect heterogeneity, since they contain "better than random noise" information.

Besides the socio-demographic variables collected in the introduction survey, we also included a number of variables from the final survey, which was conducted after the experiment was concluded. A battery of questions was used to elicit respondents' personal values (Schwartz, 1992; De Groot and Steg, 2010). Using this methodology, respondents were assigned an index along four dimensions labeled "altruistic", "egoistic", "hedonic" and "biospheric". Furthermore, we examined the extent to which the participants understood the concept of the external costs of transport that we explained to them at the beginning of the treatment, and in each of the reports (at least for the two treatment groups; no such information was provided to the control group). Specifically, we asked them to choose the definition of the external costs of transport from four possible answers.[24] We include this information as a dummy labeled "Correct EC".

We include the variables identified to be important by the CF algorithm as interaction terms with the DiD^P_{its} "post" and DiD^I_{its} dummies in (1). These variables are dummies based on gender, income, age, education, household size, language, citizenship, urbanization, car ownership, owning a half-fare public transport subscription, weekend, external cost question and the four values dimensions. To account for the correlation among these dummies, we include them jointly in a multi-variate regression.

The regression coefficients are in Table A.4. Overall, we find that the effect is relatively homogeneous across socio-demographic characteristics, with some exceptions. Setting p$\leq$ 0.05 as the threshold, we the response is stronger (i.e., more negative) for the young, those living in rural areas, car owners and those who correctly identified the definition of external transport costs. For French speakers, the effect is weaker.[25] For people that live in suburban municipalities, a higher responsiveness to information is neutralized by a lower responsiveness to pricing, such that the overall effect is similar to that of the reference category "urban".

Last, we find that the study participants that scored above the median in terms of the altruistic index responded significantly more to information alone, which is consistent with expectations. We see no differential response of altruists to pricing (suggesting that there is no"crowding out" effect). There were no statistically significant differences along the other three values-dimensions.[26]

Table 6 presents the proportional effects, price increases and resulting elasticities for the sub-samples for which we found statistically significant treatment heterogeneity. The proportional reduction in external costs for the participants that correctly identified the external costs in the "exam" question is 11.3%, whereas that for the rest of the sample is

[24]The question was formulated as "How would you define the external costs of your transport behavior?" The possible responses were: (i) the costs associated with my travel behavior that I have to pay myself; (ii) the costs imposed on society as a result of my travel behavior; (iii) the total costs associated with my travel behavior (sum of private and societal costs); (iv) I don't know what the external costs of travel are. This question was asked to all groups; the correct answer (ii) was provided by 41% in the control group, 43% in the pricing group and 46% in the information group, indicating that our explanation only had a small effect in informing people about the external costs of transport.

[25]This variable was coded based on participants' preferred language, not their region of residence. However, as most French speakers live in the French-speaking region of Switzerland, this interaction may also pick up a regional effect.

[26]Since the personal values were elicited only after the experiment, it is possible that they are influenced by the treatment. Table A.13 shows that most values are distributed equally across the treatment groups, with the exception of "biospheric" that has lower values for the pricing group. However, this variable did to turn out to be statistically significant in terms of the effect heterogeneity.

Table 6: Response for subsamples

	Treatment effect (%)			Total price increase (%)			Elasticity				
	Estimate	Lower Bound	Upper Bound	Estimate	Lower Bound	Upper Bound	Estimate	Lower Bound	Upper Bound	p	N
Total sample	-0.045	-0.073	-0.018	0.193	0.187	0.198	-0.236	-0.381	-0.091	0.001	168,362
Age>54	-0.061	-0.127	0.005	0.171	0.160	0.182	-0.356	-0.743	0.030	0.071	33,214
30≤Age≤54	-0.028	-0.062	0.006	0.193	0.185	0.201	-0.145	-0.321	0.031	0.106	91,631
Age<30	-0.071	-0.128	-0.015	0.209	0.196	0.221	-0.342	-0.611	-0.072	0.013	43,511
German	-0.063	-0.095	-0.030	0.186	0.179	0.193	-0.338	-0.512	-0.164	<0.001	112,150
French	0.001	-0.052	0.053	0.210	0.199	0.222	0.003	-0.244	0.251	0.979	43,925
English	-0.048	-0.153	0.058	0.216	0.190	0.242	-0.221	-0.714	0.272	0.380	12,280
Urban	-0.032	-0.066	0.002	0.199	0.191	0.207	-0.160	-0.332	0.012	0.068	108,174
Suburban	-0.051	-0.102	-0.001	0.184	0.175	0.194	-0.278	-0.550	-0.007	0.044	46,540
Rural	-0.112	-0.204	-0.020	0.181	0.165	0.196	-0.622	-1.129	-0.114	0.016	13,635
Car owner	-0.044	-0.074	-0.015	0.192	0.186	0.198	-0.230	-0.384	-0.077	0.003	147,616
No car	-0.054	-0.151	0.042	0.198	0.180	0.216	-0.274	-0.756	0.208	0.265	20,745
Incorrect EC	-0.002	-0.038	0.034	0.194	0.187	0.202	-0.008	-0.193	0.178	0.934	92,334
Correct EC	-0.113	-0.157	-0.069	0.191	0.181	0.200	-0.591	-0.823	-0.359	<0.001	70,843

Notes: The lower and upper bounds reflect the 95%-confidence interval, based on a bootstrap with 1,000 replications. The last columns provides the probability that the elasticity is positive and the size of the subsample. EC stands for the answer to the multiple-choice question about the definition of external costs of transport.

narrowly centered around zero. This implies that the ATE is exclusively driven by those participants that understood the concept of external costs.

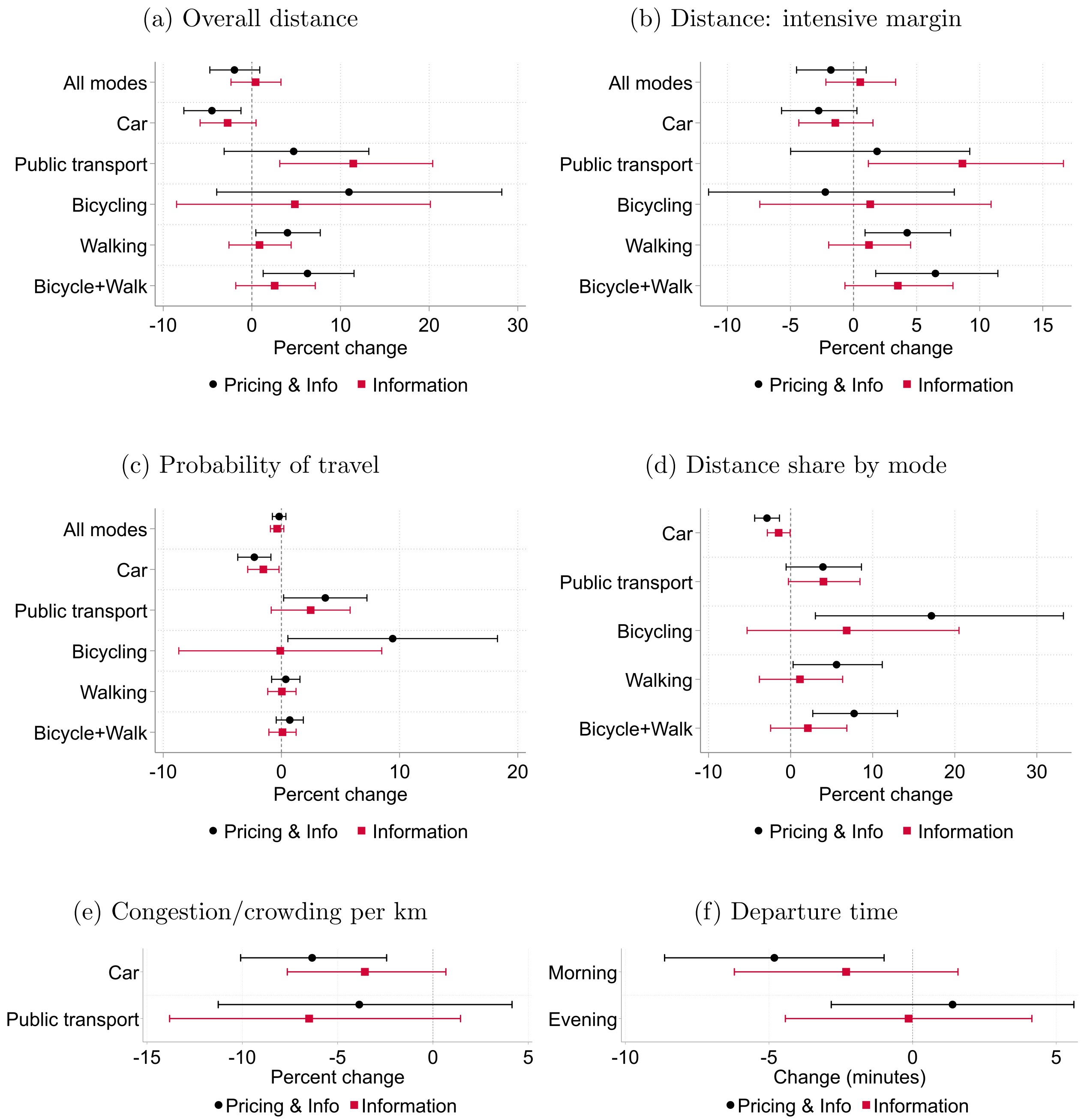
4.3 Mechanisms

People can reduce their external costs of transport in different ways. The can travel less frequently or less far, substitute towards modes associated with lower external costs or choose different routes and departure times. To shed light on potential mechanisms that mediate the reduction in external costs, Figure 6 shows the effect of the pricing treatment on various outcomes of interest (for the regression results, see Tables A.5 - A.10).

The treatment does not reduce overall travel distances for the sample as a whole, but we measure a statistically significant reduction in car distance countered by increases in the other modes. The effect can be seen separately on the intensive margin (i.e., conditional on traveling with a particular mode on a given day) and on the extensive margin (i.e., the probability of using a mode). The mode shift becomes more salient if the treatment effect is shown for mode (distance) share.

The pricing treatment significantly reduces congestion costs per km of car travel, 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

Figure 6: Mechanisms underlying the reduction in external costs



Notes: The bars denote 95%-confidence intervals. In panels (a), (b), (d) and (e), the treatment effects are computed using a Poisson pseudo-maximum likelihood (ppml) regression. Panel (c) shows the marginal results (semi-elasticity) of a logit regression. In panel (f), a linear DiD-specification is chosen with the departure time (measured in minutes after midnight) as the dependent variable. The underlying regressions are shown in Tables A.5, A.9 and A.10.

observe a statistically significant shift in the average departure time for car trips in the morning towards earlier departures, but no effect in the evening.[27] Note, however, that this departure shift is poorly identified as we cannot differentiate between a shift in car departure time from a mode shift during the same time frame, as both would affect the average car departure time measured during the observation period. There was no reduction in crowding or departure times for public transport.

As much of the effect appears to be driven by a shift away from driving, Table 7 provides the estimates for the Average Direct Effect (ADE) and the Average Indirect Effect (AIE) as defined in (6)-(7), with daily car distance as the mediator. Driving distance is confirmed as a strong mediator for the change in external costs that captures most of the ATE associated with the pricing treatment. However, there is a remaining and statistically significant ADE, which is the combination of all mediators unrelated to car-km. Given the results reported above, this unidentified effect is likely due to shifts in departure times. Interestingly, we also observe a statistically significant AIE for the information-only treatment, despite the absence of a statistically significant ATE.[28] This means that informing people about their external costs of transport significantly affects these costs via a reduction in driving, even if the total effect is not statistically significant due a combination of other responses.

Table 7: Mediation analysis (mediator: driving distance)

	Coefficient	Lower bound	Upper bound
ADE (Pricing & Information)	-0.051	-0.101	-0.003
AIE (Pricing & Information)	-0.175	-0.301	-0.048
ADE (Information only)	0.011	-0.037	0.058
AIE (Information only)	-0.107	-0.301	-0.048

Notes: The bounds show the 95%- percentile bootstrap confidence intervals, which is the recommended choice for mediation analysis (Tibbe and Montoya, 2022).

4.4 Social acceptability

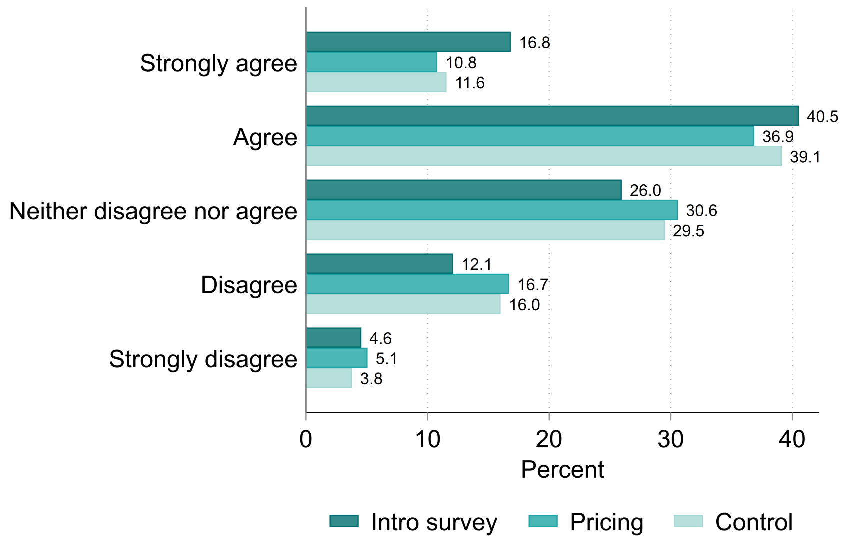
Even if transport pricing works, its implementation may be challenging not only in terms of technology and data confidentiality, but also due to social acceptability (Eliasson, 2021). To learn more about this, we asked the respondents in the final survey provide the extent to which they agreed with the following statement: "The price for mobility should reflect the

[27]For this regression, we only included people for whom we observed at least one peak-hour trip during the observation phase. We then computed the average trip departure time for trips before and after noon and used this as the dependent variable in eq. (1).

[28]The absence of a significant ATE does not preclude the existence of a significant ADE or AIE, as these could be of opposite signs. This is the case here.

social cost (e.g., health, environment, congestion)." Figure 7 shows the responses, separated by treatment group in the RCT and also for the intro survey sample. A majority of the respondents were either positive or neutral, with only 20% rejecting this statement.[29] We do not find a significant difference across the treatment groups.

Figure 7: Support for transport pricing



Notes: The figure shows participants' responses in percent to the questions described in the main text. Observations: 19,440 (intro survey),1,081 (control); 1,066 (pricing).

5 Welfare implications

The reduction in true external costs directly translates into a welfare gain, by definition. However, people experience disutility from changing their behavior in response to the introduction of taxes, and this has to be considered when making statements about welfare. To quantify the utility loss, we estimate a mode choice model on the trip level. We complement each observed trip by a set of non-chosen alternatives and compute the private and external costs associated with them. Where there is no plausible alternative (for example, no public transport available or the trip is too long to be carried out by bicycle), the choice set is adapted accordingly. We allow for variation in the estimated preference parameters by estimating a mixed logit model, which relaxes the IIA assumption in the standard multinomial

[29] We also tried two other formulations of this question. One version was worded using more technical language, with a reference to the revenue-neutral introduction of transport pricing, and another asked about the introduction of time-varying pricing but without any reference to social costs. The level of support varied across these questions. For more information, see Axhausen et al. (2021).

logit approach. More specifically, we allow the valuation of travel time to vary across people and modes, and the valuation of costs to vary across people. We further control for the weather, trip-specific purpose, departure time, as well as the first mode used on a given day, since this may preclude certain other modes from being used later in the day. As we do not include different departure times or routes, our model only captures the effect of mode choice (which is the main mediator; see above).

Most of the transport literature uses observational rather than experimental data to estimate mode choice models. The problem with this approach is that longer trips also cost more money (especially if costs are computed on a per-km-basis), such that the time and cost variables strongly co-vary. This poses a challenge to identify their separate effects. Stated choice-experiments overcome this issue by varying the the trip attributes exogenously. However, this comes at the cost of replacing observational data with stated preference data, which may not always reflect true behavior. Our experimental setting provides us with the opportunity to combine observational data with an exogenous price change, which significantly improve the credibility of the estimates relative to the two standard approaches.

We estimate the model using maximum likelihood and based on the sub-sample of respondents in the control and pricing groups that answered the question on the definition of external costs correctly.[30]

Our utility functions for the mixed logit models take the following form:

$$U_{ijt} = \beta_{0j} + \beta_{time,i} \cdot \text{Travel time}_{ijt} + \beta_{cost,i} \cdot \text{Total cost}_{ijt} + \mathbf{w}_t'\delta + \mathbf{z}_t'\gamma + \epsilon_{ijt} \quad (8)$$

$$U_{ijt} = V_{ijt} + \epsilon_{ijt}, \quad (9)$$

U_{ijt} is the utility for individual i choosing mode j for trip t. We include mode-specific constants β_0 to capture baseline utility for each mode and control for unobserved heterogeneity between modes. We allow for individual-specific, i.e. *mixed*, preferences for travel time and cost, $\beta_{time,i}$ and $\beta_{cost,i}$ and include vectors of weather-related and trip-specific variables $\mathbf{w}$ and $\mathbf{z}$. The error term ϵ_{ijt} is i.i.d. Extreme Value Type I. V_{ijt} is the observed part of utility U_{ijt}.

The panel structure of our data allows us to calculate conditional distributions for the random parameters in the mixed logit models. These are generally tighter than the unconditional parameter distributions and allow us to approximate an individual's position within the unconditional distribution. Table A.11 presents the estimated coefficients from the preferred mode choice model.

[30]Including the people in the information group, or the respondents in the pricing group that were not engaged in the experiment (and thus showed now response at all) would just add noise and likely lead to a downward bias in the cost sensitivity.

Table 8: Value of travel time savings and price elasticities

VTTS (CHF/h)	Mean	Median	SD	Min	Max
Car	1.97	0.37	10.43	0.13	189.95
PT	10.96	10.96	0.00	10.89	10.99
Bike	10.91	10.91	0.03	10.70	11.41
Walk	67.47	55.59	55.97	3.53	472.13

Price elasticities:	Choice probabilities				Distance				Travel time			
	Car	PT	Bike	Walk	Car	PT	Bike	Walk	Car	PT	Bike	Walk
Price for driving	-0.17	0.20	0.07	0.02	-0.45	0.45	0.11	0.03	-0.35	0.36	0.11	0.03
Price for PT	0.07	-0.15	0.03	0.01	0.18	-0.25	0.05	0.02	0.13	-0.23	0.05	0.02

Notes: The VTTS values represent the conditional distributions of willingness to pay to save one hour of travel time with the respective mode. For the elasticities, the row refers to the mode for which the price has been increased by 1% and the column refers to the percent change in choice probability, distance, and travel time, respectively.

Dividing the coefficients on time by the cost parameter yields an estimate for the value of travel time savings (VTTS). Table 8 presents the estimates of the conditional VTTS for each mode, as well as the own- and cross-price elasticities for car and public transport for the mode choice probability, distance, and travel time, respectively. The low VTTS for car is likely due to the car use participation criterion for MOBIS, such that our participants are not averse to car travel. We do not interpret the VTTS values further other than to say that they are within reasonable bounds, since our experiment was not designed to uncover "true" VTTS values for Switzerland. The own price elasticities show that car kilometers and travel time are more elastic than for PT, with the opposite the case for the cross price elasticities. In fact, the cross price elasticities for car mode choice, distance, and travel time suggest that reductions in car choice and demand are more than offset by increases in PT, bike, and walking choice and demand. For an increase in PT cost, the reduction in PT choice and travel time is not fully offset by car choice and demand, whereas it is just offset for distance.

To compute the utility loss from the pricing treatment, we use the result from McFadden (1977) and Small and Rosen (1981), and write the expected consumer surplus from each trip for each individual

$$E\left[CS_{it}\right] = E\left[\frac{1}{|\beta_{cost,i}|}\max_{j\in J} U_{ijt}\right], \tag{10}$$

where U_{it} is the utility for individual i for trip t. Small and Rosen (1981) show that even if utility U_{it} is not observable, expected consumer surplus can be calculated via observed utility V_{it} if the error terms are distributed i.i.d. Extreme Value Type I as in (8) and (9). Dividing the expected utility by the marginal utility of income $\beta_{cost,i}$ yields a money-metric value of consumer surplus. We can rewrite eq. (10) as

$$E\left[CS_{it}\right] = \frac{1}{|\beta_{cost,i}|} \cdot \ln\left(\sum_{j=1}^{J} \exp(V_{ijt})\right) + C, \tag{11}$$

which is known as the *logsum* for individual i and trip t.[31]

Leveraging our experiment, we can calculate not only the difference in utility before and after a change in attribute values, but a difference in differences, which controls for any temporal trends that may exist in the data, and also allows us to capture the change in travel distance in response to the treatment.[32] We thus use expression (11) as the dependent variable in the difference-in-differences model in eq. (1) that was used to derive the reduced-form results. This procedure allows us to control for unobserved heterogeneity in monetized utility space, which is much easier than including person, calendar day and day-of-study fixed effects in a mixed logit model.

We calculate the tax revenue and the reduction in external costs within the model.[33]

Table 9: Welfare effects from Pigovian pricing and alternative policies

	Pigovian pricing	Perimeter pricing	Fuel tax
Monetised utility loss	-2.893	-2.567	-2.733
Average payment	2.697	2.697	2.697
Reduction in external costs	0.579	0.203	0.228
Welfare gain	0.383	0.333	0.192
Welfare gain (per person and year)	139.80	121.55	70.08
as share of Pigovian welfare gain	100%	86.9%	50.1%
N	41,143	41,143	41,143

Note: Values in CHF per person-day (unless otherwise stated).

Table 9 presents the results. The average daily monetized utility loss is CHF 2.89, which is a 3% decrease relative to the monetized utility in the absence of pricing. Subtracting the deadweight loss of taxation from the gain due to the reduction in external costs leads to an

[31] Since the change in costs faced by the respondents is small compared to total income, we can abstract from income effects when moving from utility to monetary space.

[32] For the subsample that correctly identified the external costs of transport, which is the basis for this structural model, the treatment reduced the overall travel distance by 7%. By using the actually recorded trips in a DiD framework, our analysis captures this change in overall travel demand, along with the change in mode.

[33] Since our sample for this analysis departs from the whole sample, we cannot use the tax payment and the change in external costs from the reduced-form analysis but have to compute them anew.

annual welfare gain of around CHF 140 per person.[34]

Note also that the conceptual model that underlies this exercise is that the entire tax revenue is redistributed lump-sum. If, however, the revenue were returned in a progressive fashion, additional distributional gains could be obtained. We abstain from including distributional gains in this section as they depend not only on the distribution of revenue recycling but also on the way society trades off the utilities of different people (i.e., the elusive social welfare function). Note also that when asked about the preferred use for the revenue from transport pricing, the majority of the participants indicated that at least some of the money to be used for transport projects, whereas fewer than 8% preferred the revenue to be returned to households (see Fig. A.1).

We can further use our model to compute the welfare implications of instituting different versions of second-best pricing. We start with example a fuel tax. To obtain a meaningful comparison, we calibrate the alternative policy such that the the resulting revenue, taking into account the behavioral responses, is the same as with the Pigovian tax (alternatively, we could have conditioned on the reduction of external costs). This requires simulating the model using a series of different tax levels, until the revenue is equal to the the target. As shown in Table 9, the resulting welfare gain is about CHF 70 per person and year, which is just over half of the Pigovian pricing welfare gain.

Second, we compute the welfare effects of a congestion tax levied within urban areas during peak hours, and which is levied both on drivers and public transport users. This policy is only slightly less complex than our Pigovian approach as it levies taxes that vary over time, space and mode. We find that such a tax would result in a welfare gain of CHF 122 per person and year, or 87% of the Pigovian welfare gain.

We stress that these calculations derive an upper bound for the welfare effects from alternative policies, as our model can only capture the effect on mode choice while holding the number of trips, their timing and the total distance fixed. If a fuel tax or perimeter pricing policy also changes the travel volume and/or shifts demand within the day, the substitution effect would be larger and thus the welfare gain smaller.[35]

[34]This is about four times higher than the welfare estimates computed by Almagro et al. (2024) for an optimal transport policy in Chicago, but that analysis excludes some of the external costs considered here (e.g., the health costs related to accidents or noise).

[35]The welfare gain is the change in external costs net of the deadweight loss, defined as the difference between the monetized utility loss and the tax revenue. As is well-known from standard tax theory, the deadweight loss increases with the substitution effect. By holding the number of trips constant, we provide a lower bound for the deadweight loss and thus an upper bound for the welfare gain from an alternative policy.

6 Internal and external validity

In this section, we discuss the main threats to identification and the extent to which we believe that our results may hold also in other settings.

6.1 Strategic app manipulation

Participants were invited to use the validation interface to confirm the detected mode and purpose of their trips and activities. And they made use of this function: On average, we record 0.32 corrections per person-day, and at least one correction on 18% of the person-days. As the mode is crucial in determining the external costs, the possibility of overwriting the detected mode for a particular trip leg 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, manual mode adjustments could also be truthful corrections of a mis-assigned mode. The key question is whether we observe systematically different mode correction behavior for the treatment groups relative to the control group. To test for this, we use the mode corrections as the dependent variable in our difference-in-differences regression. Columns (1)-(2) in Table 10 show that there is no difference in the number of corrections per day across groups, nor in the probability of at least one correction taking place per person-day.[36]

Table 10: Mode corrections and spatial jumps

Dependent variable	Nr. of corrections	Correction (1/0)	Nr. of jumps	Jump (1/0)
Pricing	-0.042	-0.006	0.059	0.003
	(0.042)	(0.009)	(0.073)	(0.004)
Information	-0.056	-0.012	-0.076	-0.004
	(0.037)	(0.008)	(0.063)	(0.004)
N	112,744	112,744	107,993	107,993

Notes: Standard errors (in parentheses) clustered at participant level. The dependent variable in cols. (1) & (3) is the number of mode corrections and spatial jumps per day, respectively. The coefficients are proportional effects, estimated using a ppml model. Cols. (2) & (4) display the marginal effects from logit regressions on dummies denoting whether at least one correction or jump was recorded on a given day. All regressions control for person, calendar and study day FE.

Another form of potential manipulation would consist in participants turning off the app before departure and switching it back on once they have reached their destination. To

[36]Note that these tests are done only with people who recorded at least one correction or one spatial jump, as the pure zeros are perfectly identified by the person-FE. This is the reason for the smaller sample relative to the other regressions.

investigate this, we marked all instances in which we see "spatial jumps" in the data, in the sense that a participant's location at the end of one trip is not the same as the starting location of the following trip. To abstract from small random jumps (due to, e.g., cellphone reception gaps or brief pauses in the GPS signal), we set a limit of 10km for a significant spatial jump.[37] We record an average of 0.04 jumps per person-day. Columns 3-4 in Table 10 show that there is no effect of the treatment indicator on the number and probability of such spatial jumps. 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. The resulting treatment effects are shown in column 2 of Table 11. The ATE is effectively unchanged.

Last, we can compare the distances by mode with and without the correction (note that we did not compute the external costs associated with the originally detected but later corrected trip stages). In the appendix we show that these distances are very similar (Table A.6), and that the ATE on distance by model is essentially unchanged (Tables A.5 vs. A.7). Taken together, these robustness tests imply that our results are unlikely to be driven by strategic mode correction.

Table 11: Subsample analyses

	Baseline	w/o corrections	w/o weeks 7-8	w/o weeks 5-6	w/o zeros
Pricing	-0.226**	-0.226**	-0.243**	-0.242**	-0.224**
	(0.070)	(0.077)	(0.080)	(0.083)	(0.072)
Information	-0.096	-0.093	-0.088	-0.100	-0.092
	(0.067)	(0.074)	(0.079)	(0.081)	(0.069)
Adj. R^2	0.234	0.238	0.237	0.233	0.239
Clusters	3,616	3,616	3,616	3,616	3,616
N	168,362	137,771	126,060	128,156	160,974

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

6.2 Missing tracking data

Although tracking discipline was extremely high in our sample, many participants did not record positive distances on all days. To differentiate between true zeros (i.e., participants

[37]The results do not change when we use a threshold of 5km or 20km instead.

no traveling) and missings (participants disabling the app or traveling without their mobile phone), we rely on imputed activities to link stages. Suppose that a participant travels home on Friday evening and does not deliver another track until Monday. If the app imputes an uninterrupted activity (in this case labeled as "at home") that lasts from Friday to Monday, then we assign a travel distance of zero for Saturday and Sunday. However, there is not always an uninterrupted activity between stages on different days. For example, if a participant disables the app or leaves Switzerland during the study, there will not be a continuous activity linking stages, and as a consequence we treat such person-days as missings rather than zeros. The imputation of activities and locations is not always correct. To gauge the sensitivity of our results to the distinction between zeros vs. missing data, we re-estimate the model using only data from days with positive travel distances. The resulting ATE is shown in column 3 of Table 11. As it is very similar to the baseline, we believe that the ATE is unlikely to be driven by missing data mistakenly coded as zero (or vice versa).

6.3 Attrition and observability

The assignment into groups was randomized, but people could drop out of the study or switch off the app at any time. Our incentive payment of CHF 100 was paid only at the end of the study and thus designed to keep attrition low. We excluded people from the study who did not track on at least 11 days during the observation period, but this does not pose a challenge for internal validity as it occurred before the treatment assignment. What we worry about, however, is nonrandom observability during the treatment phase. If observability is correlated with the treatment assignment, or with factors that co-determine the treatment effect, then our estimate of the ATE could be biased.

A bias due to observability cannot be directly tested, but we engage in two indirect tests to examine this possibility. First, if attrition in the sense of participants permanently turning off the app were influenced by the group assignment, any bias due nonrandom attrition 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) and the last two weeks of the treatment period (column 4). The results remain largely unchanged.

Second, given our panel setting, people may not attrit completely but vary the degree to which they are observed on any given study day. If observability is related to the treatment effect, it could bias our results even if it is balanced across treatment groups. To investigate this, we regress the number of observed tracking days (ranging from 0 to 27) on pre-treatment external costs and the pre-determined characteristics over which we found the treatment effect to vary (see subsection 4.2).[38] The results show that observability correlates positively with pre-treatment external costs (column 1 in Table A.14). However, when conditioning on pre-treatment observability (as we do implicitly in our alysis by including person fixed-effects), this effect vanishes (column 2). Furthermore, the ATE does not systematically vary over pre-treatment observability (column 3). To summarize, conditional on pre-treatment characteristics, participants' observability during the treatment period is not systematically related to variables that co-determine the treatment effect. Based on these findings, we conclude that our results are not biased by nonrandom attrition / observability.

6.4 Mean reversion

People with high external costs during the observation period reduced their external costs more during the treatment period than people with low external costs. The presence of mean reversion can be seen in the first column of Table A.12 in the Appendix, where we regress the external costs of the control group on an indicator for the treatment period and interaction terms of this indicator with pre-treatment cost quartiles.

Mean reversion was expected to be present in the data to some extent, and it was an important reason for assigning personal budgets that exceeded pre-treatment external costs by 20%. As long as mean reversion affects the treatment and control groups equally, it should have no effect on the ATE. This cannot be tested directly, but to provide some indication that our treatment effect is not caused by mean reversion, we estimate our base regression separately by cost quartile (columns 2-4 in Table A.12). Conditioning on pre-observation costs means that we compare people that exhibit similar pre-treatment shocks and thus have similar tendencies for mean reversion. The results indicate that the treatment effect persists for the top three cost quartiles. For the bottom cost quartile, there is no significant effect, presumably because this group finds it difficult to further reduce their external costs of transport.

6.5 External validity

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

[38]Because we removed study day 29, the maximum number of observed days during the treatment period is 27. Note that for this analysis, we also include the 71 participants that did not deliver any observations during the treatment phase; see section 3.1.

As we argue above, attrition/observability is not determined by variables that moderate the treatment effect. And since our experiment did not introduce new tasks but simply observed people in their everyday travel, the naturalness condition is arguably not a problem here. In the following, we will therefore focus on selection and scaling.

Selection

Our sample is quite similar to the general population living in Swiss urban agglomerations in terms of socio-demographic characteristics (see section 2.1), such that one may be tempted to conclude that the results generalize directly. However, due to self-selection into the study itself, it is possible that our sample differs from the target population in terms of unobservable characteristics that are related both to the decision to participate and personal transport choices. For this reason, we cannot guarantee external validity given our sample selection procedure even when conditioning on observables.

We were careful not to make any reference to transport pricing or external costs when inviting people to participate in the study. In order for our results to mis-represent the response of the general population, there would need to be a correlation between the propensity to participate in a loosely defined “transport study” (not a Pigovian pricing experiment) and the extent to which someone responds to information and pricing associated with the external costs of transport. The fact that our treatment effects are homogeneous across most socio-economic characteristics suggests (but by no means proves) that this may not be a large source of bias. One way to get more information about the bias that arises from the self-selection problem would have been to randomize the incentive payment. Unfortunately, this was not done here as everyone was offered CHF 100 for participation.

Self-selection into the study is clearly problematic if the goal is to predict the effects of instituting Pigovian transport pricing as government policy. However, the implementation could also take other forms. For example, if faced with political opposition due to privacy concerns, transport pricing could be offered to volunteers who, in exchange, are exempt from vehicle registration taxes (or receive some other compensation via the tax code). In such an implementation, the target population could be quite similar to our sample, such that self-selection in our study would become a feature rather than a source of bias.

Scaling

We differentiate between horizontal scaling (application of our results to other populations) as well as longitudinal scaling. We start with the former.

Due to the richness of our data, we can compute conditional average treatment effects and thus predict the likely response of target populations that have a different distribution of underlying characteristics than our sample, as long as there is some common support. For example, even though our study focused on urban agglomerations, there are nevertheless a significant number of participants that live in municipalities that can be described as “rural”, such that expected treatment effects can be computed also for areas outside of cities. In cases where linearity may be assumed, it is even possible to extrapolate outside our support. Our study includes only people who drive on at least two days per week, but since we have a wide distribution of driving, one could approximate the effect for those that drive only 1.5 times per week, on average. This becomes more problematic for characteristics for which there is no common support between our sample and the target population. For example, our study did not include people outside the age range of 18-65, such that we cannot make valid predictions about the response of pensioners or children.

The majority of the treatment effect is due to mode substitution away from driving, which requires that public transport be available. It is obvious that that not all populations have access to public transport that is comparable to the Swiss setting. However, our setting is by no means unique. For example, the mode share for the city of Chicago (Almagro et al., 2024) is very similar to that in our experiment.

We would further expect general equilibrium effects such as increased travel speeds during peak hours to materialize as transport pricing is scaled to a larger portion of the population. This would increase the utility of those travelers that are not willing or able to shift outside of peak time (which is currently not considered in our partial-equilibrium welfare analysis). Conversely, these same equilibrium effects may reduce some of the response as driving during peak becomes more attractive.

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). Conversely, there is evidence that people treat “house (gambling) money” differently from “real” money. Thaler and Johnson (1990) show that individuals tend to combine prior gains (in this study the accumulated budget over the 4 week observation phase) with subsequent losses, which, as long as they are less than the initial gain, are seen as a “non-gain” rather than a loss. This facilitates risk-seeking behavior until the prior gain is completely depleted. Since our participants start with a gain (i.e., their personal budget), this would lead to an under-estimate of the effect relative to transport pricing that would become part of households’ general expenditure.

Scaling also concerns the time frame of the experiment. The experiment took place in the months of September through January in Switzerland. Although this includes a number of weeks with relatively mild climate, the colder part of the year clearly dominates. To the extent that cycling and walking (and, by extension, using public transport, which usually requires some access on foot) are more attractive in the warmer months, our experiment may under-estimate the effect over the whole year.

Most importantly, the treatment period in our experiment lasted only one month, such that we can only measure short-term responses. With a permanent introduction of transport pricing, additional margins of response will become available such as the choice of work and home locations, changes in activity routines, vehicle/transit pass ownership or negotiations with employers about work hours and location. Studies of fuel-price elasticities indicate that the long-term response is about twice as high as the short-term response (Goodwin et al., 2004). Furthermore, the behavioral response was concentrated among those respondents who understood the concept of external costs underlying the pricing. Whereas it is to be expected that not everyone pays close attention to the "rules of the game" in a short study, a general introduction of transport pricing would presumably have a greater salience and people would learn over time how to avoid costs. On the other hand, it is conceivable that the salience of the treatment, and in particular the effect of providing information, may fade over time, and people may discover better how to cheat the system (depending on how this would be implemented). Both of these adjustments would reduce the long-term impact of this policy. However, we would expect the effect via the greater elasticity in the long run to dominate. Future studies are needed to better understand the implications of externalities-based transport pricing in the long run.

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

7 Conclusion

The MOBIS experiment implemented transport pricing based on the social marginal costs. The short-term response for total external costs associated with the pricing treatment was -4.5%. Whereas the information-only treatment had a strong effect for subgroups of the population (such as altruists), the effect is only marginally statistically significant for the sample overall. However, our results imply that both information and monetary incentives play an important role in explaining the behavioral change in our experiment. The reduction in the external costs is due to a combination of a shift away from driving towards other modes and towards less congested times and routes. The effect varies with age, degree of urbanity, car ownership and language region, and particularly strongly with the degree to which participants engaged with the experiment. The effect is entirely driven y those that understood the concept of external costs.

MOBIS is the first multi-modal RCT investigating Pigovian pricing in a transport setting and thus different to uni-modal pricing schemes. We therefore cannot say whether our estimates are large or small. As it happens, the elasticity of -0.24 that we recover is in the same range as estimates of the short-run fuel elasticities (Goodwin et al., 2004) and results based on toll pricing (Bain, 2019). On the other hand, our proportional response is smaller than that measured by Nielsen (2004) and Leape (2006), most likely because these studies use pricing interventions that exceed the Pigovian rate.

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

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 level in the political process and to coordinate between different levels of government (e.g., cities vs. regions; see Eliasson, 2021). 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. 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.

Almagro, Milena, Felipe Barbieri, Juan Camilo Castillo, Nathaniel G Hickok and Tobias Salz (2024). "Optimal Urban Transportation Policy: Evidence from Chicago."Technical report, National Bureau of Economic Research Working Paper Nr. 32185.

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

Axhausen, Kay W., J. Molloy, C. Tchervenkov, F. Becker, B. Hintermann, B. Schoeman, T. Götschi, A. Castro and U. Tomic (2021). "Empirical Analysis of Mobility Behavior in the Presence of Pigovian Transport Pricing."Technical report, ASTRA. Report Nr. 1704.

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

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

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

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

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.

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.

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." in *16th Swiss Transport Research Conference (STRC 2016)*. 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).

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

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." in *Proceedings of the 6th ACM workshop on Next generation mobile computing for dynamic personalised travel planning.*: 9–10.

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

Christensen, Peter and Adam Osman (2023). "The demand for mobility: Evidence from an experiment with uber riders."Technical report, National Bureau of Economic Research.

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, p. 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."Technical report, European Environment Agency. Indicator Assessment.

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

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

Federal Roads Office (2017). "Handbook NISTRA 2017." technical report, Federal Roads Office, Bern. `https://www.astra.admin.ch/astra/de/home/fachleute/dokumente-nationalstrassen/fachdokumente/nistra.html`.

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

——— (2022). "Kosten und Finanzierung des Verkehrs 2019 (Costs and Financing of Transport in 2019)."Technical report. https://www.bfs.admin.ch/asset/de/811-1900.

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

Goldszmidt, Ariel, John A List, Robert D Metcalfe, Ian Muir, V Kerry Smith and Jenny Wang (2020). "The value of time in the United States: Estimates from nationwide natural field experiments."Technical report, National Bureau of Economic Research.

Goodwin, Phil, Joyce Dargay and Mark Hanly (2004). "Elasticities of road traffic and fuel consumption with respect to price and income: a review." *Trans. 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.

Hintermann, Beat, Beaumont Schoeman, Joseph Molloy, Thomas Schatzmann, Christopher Tchervenkov and Kay W Axhausen (2023). "The impact of COVID-19 on mobility choices in Switzerland." *Transportation Research Part A: Policy and Practice* 169, p. 103582.

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

IEA (2020). "Tracking Transport 2020."Technical report, International Energy Agency.

Imai, Kosuke, Luke Keele and Teppei Yamamoto (2010). "Identification, inference and sensitivity analysis for causal mediation effects."

INRIX (2020). "Global Traffic Scorecard 2019."Technical report, Trevor Reed.

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

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

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

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

Kraemer, Helena Chmura, Michaela Kiernan, Marilyn Essex and David J Kupfer (2008). “How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches..” *Health Psychology* 27(2S), p. S101.

Kreindler, Gabriel (2023). “Peak-hour road congestion pricing: Experimental evidence and equilibrium implications.” *NBER Working Paper Nr. 30903*.

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

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

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

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

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

McFadden, Daniel (1977). “Modelling the Choice of Residential Location.” *Cowles Foundation Discussion Papers* 710.

Molloy, Joseph, Alberto Castro, Thomas Götschi, Beaumont Schoeman, Christopher Tchervenkov, Uros Tomic, Beat Hintermann and Kay W Axhausen (2023). “The MOBIS dataset: a large GPS dataset of mobility behaviour in Switzerland.” *Transportation* 50(5): 1983–2007.

Molloy, Joseph, Thomas Schatzmann, Beaumont Schoeman, Christopher Tchervenkov, Beat Hintermann and Kay W Axhausen (2021). “Observed impacts of the Covid-19 first wave on travel behaviour in Switzerland based on a large GPS panel.” *Transport Policy* 104: 43–51.

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, p. 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." in *17th Swiss Transport Research Conference*.

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*: 31–39.

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

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

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

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

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*. 25 Elsevier: 1–65.

Small, K. A. (2008). "Urban transportation policy: A guide and road map." in *Unraveling the Urban Enigma: City Prospects, City Policies, Conference and Book*. Wharton School, University of Pennsylvania.

Small, Kenneth A and Harvey S Rosen (1981). "Applied welfare economics with discrete choice models." *Econometrica: Journal of the Econometric Society*: 105–130.

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."Technical report. `https://www.are.admin.ch/are/en/home/mobility/data/mtmc.html`.

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

Tibbe, Tristan D and Amanda K Montoya (2022). "Correcting the bias correction for the bootstrap confidence interval in mediation analysis." *Frontiers in psychology* 13, p. 810258.

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

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.

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

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.

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.

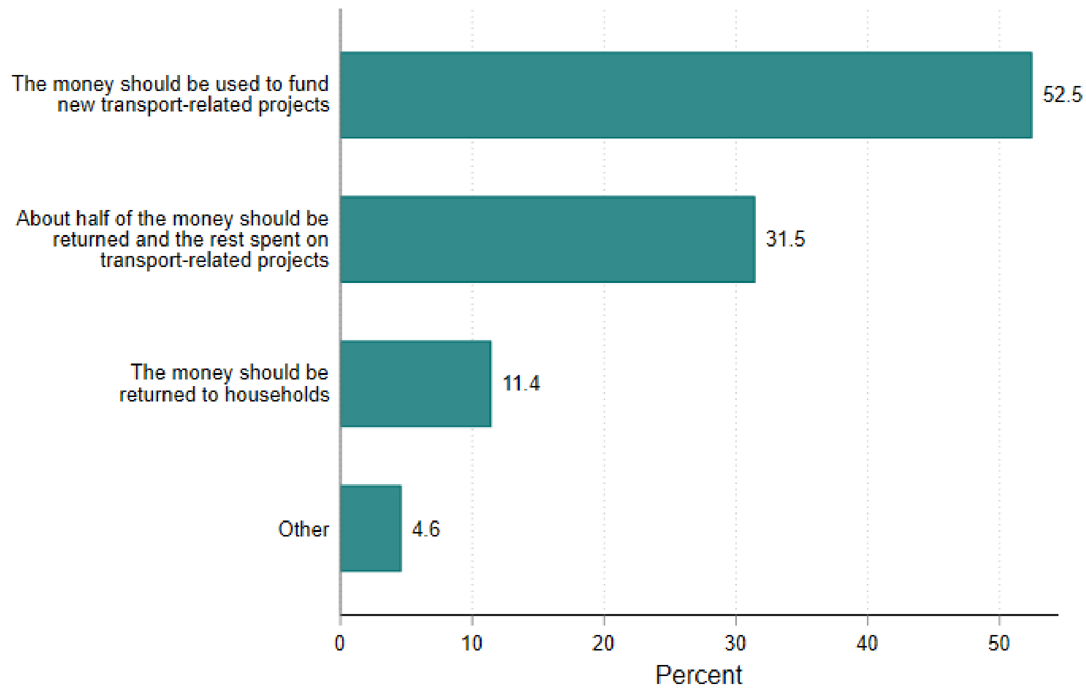
Wu, Linlin, Biao Yang and Peng Jing (2016). "Travel mode detection based on GPS raw data collected by smartphones: A systematic review of the existing methodologies." *Information* 7(4), p. 67.

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

Online Appendix

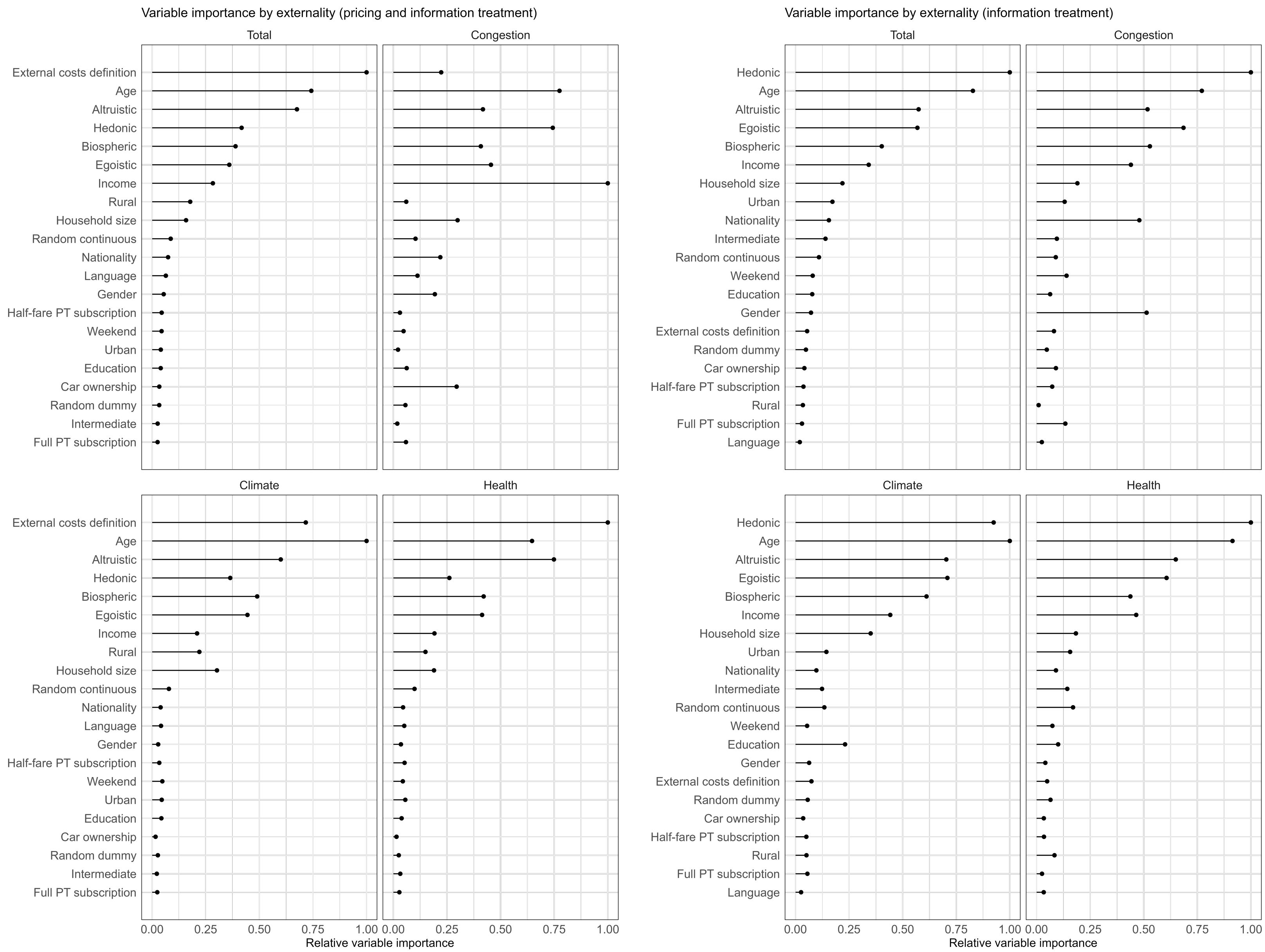
A Additional tables and figures

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?". Observations: N(overall)=3418.

Figure A.2: Variable importance in Causal Forest



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

Table A.1: Tracking summary statistics, by mode

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

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

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

Table A.2: Distribution of external costs across modes and cost dimensions

	CO2	Pollution	Accidents	Physical act.	Congestion	Crowding	Total
Car	19.17%	42.31%	15.89%		19.19%		96.6%
Train	0.01%	1.34%	0.10%			1.69%	3.1%
Light rail		0.74%	0.04%			0.68%	1.5%
Tram		0.02%	0.13%			0.22%	0.4%
Bus	0.34%	0.88%	0.33%			0.28%	1.8%
Bicycle			3.71%	-2.71%			1.0%
Walk			2.95%	-7.32%			-4.4%
Total	19.52%	45.30%	23.15%	-10.03%	19.19%	2.87%	100.0%

Notes: The table shows a heat map of the relative contribution of each cell to the total external costs observed in the sample (both phases, all groups). Due to the benefits from physical activity, the positive costs add up to more than 100%.

Table A.3: Treatment effects for car and public transport

(a) Car

	Total External Costs	Health Costs	Climate Costs	Congestion Costs
Pricing	-0.232**	-0.124**	-0.040*	-0.068**
	(0.070)	(0.044)	(0.016)	(0.020)
Information	-0.116'	-0.067	-0.022	-0.027
	(0.067)	(0.042)	(0.015)	(0.020)
Difference	-0.116'	-0.057	-0.017	-0.041*
	(0.070)	(0.044)	(0.016)	(0.020)
Adj. R^2	0.240	0.231	0.226	0.271
Clusters	3,616	3,616	3,616	3,616
N	168,362	168,362	168,362	168,362

(b) Public Transport

	Total External Costs	Health Costs	Climate Costs	Congestion Costs
Pricing	0.008	0.009'	0.001	-0.002
	(0.013)	(0.005)	(0.001)	(0.010)
Information	0.019	0.014**	0.001	0.004
	(0.012)	(0.005)	(0.001)	(0.009)
Difference	-0.012	-0.005	0.000	-0.007
	(0.012)	(0.005)	(0.001)	(0.010)
Adj. R^2	0.285	0.242	0.172	0.269
Clusters	3,616	3,616	3,616	3,616
N	168,362	168,362	168,362	168,362

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Standard errors in parentheses and clustered at the 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.4: Multivariate interactions

	Total Costs		Health Costs		Climate Costs		Congestion Costs	
	Pricing	Info.	Pricing	Info.	Pricing	Info.	Pricing	Info.
Base	0.627**	0.133	0.289'	-0.001	0.112*	0.042	0.227**	0.091
	(0.241)	(0.237)	(0.150)	(0.151)	(0.056)	(0.055)	(0.071)	(0.072)
Male	-0.117	-0.153	-0.025	-0.019	-0.006	-0.019	-0.085**	-0.116**
	(0.106)	(0.098)	(0.066)	(0.061)	(0.024)	(0.023)	(0.032)	(0.030)
Income>12k	-0.122	0.021	-0.013	0.030	-0.012	-0.015	-0.097*	0.007
	(0.131)	(0.131)	(0.079)	(0.082)	(0.032)	(0.030)	(0.045)	(0.041)
Income<8k	0.202'	-0.028	0.079	-0.060	0.042	-0.006	0.081*	0.039
	(0.120)	(0.114)	(0.075)	(0.071)	(0.027)	(0.026)	(0.036)	(0.034)
Age>54	-0.183	0.015	-0.113	-0.016	-0.061'	0.001	-0.008	0.030
	(0.142)	(0.145)	(0.089)	(0.089)	(0.033)	(0.034)	(0.044)	(0.040)
Age<30	-0.489**	0.135	-0.252**	0.082	-0.109**	0.025	-0.128**	0.028
	(0.146)	(0.116)	(0.089)	(0.074)	(0.034)	(0.027)	(0.045)	(0.037)
Tertiary ed.	-0.076	-0.034	-0.080	-0.035	-0.017	-0.020	0.021	0.021
	(0.113)	(0.110)	(0.068)	(0.070)	(0.026)	(0.026)	(0.038)	(0.031)
HH size>4	-0.205	0.192	-0.143	0.110	-0.085*	0.036	0.023	0.046
	(0.176)	(0.191)	(0.110)	(0.109)	(0.042)	(0.050)	(0.047)	(0.056)
HH size<3	-0.130	0.013	-0.013	0.026	-0.032	-0.018	-0.085*	0.005
	(0.118)	(0.112)	(0.072)	(0.070)	(0.027)	(0.026)	(0.037)	(0.033)
French sp.	0.275*	0.037	0.153*	0.006	0.058*	-0.004	0.064'	0.035
	(0.119)	(0.114)	(0.077)	(0.073)	(0.027)	(0.027)	(0.035)	(0.033)
English sp	0.338'	0.057	0.134	0.051	0.069	0.002	0.134*	0.004
	(0.202)	(0.224)	(0.120)	(0.144)	(0.049)	(0.052)	(0.066)	(0.067)
Foreign	-0.259'	0.150	-0.116	0.070	-0.045	0.040	-0.099*	0.040
	(0.137)	(0.131)	(0.082)	(0.084)	(0.032)	(0.030)	(0.045)	(0.040)
Suburban	-0.010	0.261*	-0.016	0.143*	-0.007	0.045'	0.013	0.073*
	(0.118)	(0.109)	(0.074)	(0.069)	(0.027)	(0.025)	(0.033)	(0.032)
Rural	-0.436*	0.112	-0.269*	0.055	-0.118*	0.014	-0.048	0.043
	(0.204)	(0.178)	(0.130)	(0.124)	(0.048)	(0.042)	(0.057)	(0.040)
Car owner	-0.352*	-0.113	-0.126	-0.005	-0.048	-0.015	-0.179**	-0.093'
	(0.170)	(0.158)	(0.104)	(0.096)	(0.039)	(0.037)	(0.051)	(0.051)
1/2 fare	-0.103	-0.036	-0.077	-0.009	-0.025	0.005	-0.000	-0.032
	(0.104)	(0.101)	(0.065)	(0.064)	(0.024)	(0.023)	(0.031)	(0.030)
Weekend	0.145	0.211'	0.074	0.118	0.045'	0.029	0.025	0.064'
	(0.121)	(0.117)	(0.076)	(0.074)	(0.026)	(0.025)	(0.034)	(0.036)
Correct EC	-0.363**	-0.108	-0.230**	-0.081	-0.079**	-0.039'	-0.054	0.012
	(0.109)	(0.100)	(0.068)	(0.062)	(0.025)	(0.023)	(0.033)	(0.031)
Egoistic	0.041	-0.097	0.048	-0.075	0.025	-0.042'	-0.032	0.019
	(0.105)	(0.108)	(0.064)	(0.068)	(0.024)	(0.025)	(0.033)	(0.031)
Altruistic	-0.154	-0.245*	-0.092	-0.133'	-0.032	-0.049*	-0.030	-0.064*
	(0.115)	(0.105)	(0.072)	(0.068)	(0.026)	(0.024)	(0.036)	(0.031)
Hedonic	0.066	-0.109	0.022	-0.015	0.009	-0.005	0.034	-0.089**
	(0.113)	(0.107)	(0.069)	(0.067)	(0.026)	(0.025)	(0.036)	(0.031)
Biospheric	-0.158	0.112	-0.087	0.065	-0.035	0.029	-0.036	0.018
	(0.118)	(0.112)	(0.074)	(0.072)	(0.027)	(0.025)	(0.036)	(0.032)
Adj. R^2	0.234		0.227		0.224		0.266	
Clusters	3,486		3,486		3,486		3,486	
N	163,182		163,182		163,182		163,182	

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

Table A.5: ATE on travel distance

(a) Overall margin

	Total	Car	PT	Bicycle	Walking	Slow
Pricing	0.980	0.955**	1.047	1.110	1.040*	1.063*
	(0.014)	(0.016)	(0.042)	(0.082)	(0.019)	(0.026)
Information	1.004	0.973'	1.114**	1.049	1.009	1.026
	(0.014)	(0.016)	(0.044)	(0.073)	(0.018)	(0.023)
Difference	0.976'	0.982	0.940'	1.058	1.031'	1.036
	(0.014)	(0.017)	(0.034)	(0.081)	(0.018)	(0.026)
Adj. R^2	0.270	0.297	0.413	0.469	0.266	0.348
Clusters	3,616	3,615	3,496	2,210	3,616	3,616
N	168,362	168,350	162,874	103,604	168,362	168,362

(b) Intensive margin

	Total	Car	PT	Bicycle	Walking	Slow
Pricing	0.982	0.972'	1.019	0.978	1.042*	1.065**
	(0.014)	(0.015)	(0.036)	(0.050)	(0.017)	(0.025)
Information	1.005	0.986	1.086*	1.013	1.012	1.035
	(0.014)	(0.015)	(0.039)	(0.047)	(0.017)	(0.022)
Difference	0.977'	0.987	0.938'	0.965	1.030'	1.029
	(0.013)	(0.015)	(0.032)	(0.053)	(0.017)	(0.024)
Adj. R^2	0.277	0.296	0.483	0.642	0.260	0.366
Cluster	3,616	3,614	3,339	1,589	3,615	3,615
N	160,974	131,081	56,493	13,432	138,572	140,452

(c) Extensive margin

	Total	Car	PT	Bicycle	Walking	Slow
Pricing	0.998	0.977**	1.038*	1.099*	1.004	1.007
	(0.003)	(0.007)	(0.019)	(0.050)	(0.006)	(0.006)
Information	0.996	0.985*	1.025	0.999	1.000	1.001
	(0.003)	(0.007)	(0.017)	(0.044)	(0.006)	(0.006)
Difference	1.002	0.992	1.012	1.100*	1.003	1.006
	(0.003)	(0.007)	(0.018)	(0.052)	(0.006)	(0.006)
Adj. R^2	0.002	0.029	0.142	0.201	0.017	0.016
Clusters	3,616	3,615	3,496	2,210	3,616	3,616
N	168,362	168,350	162,874	103,604	168,362	168,362

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. Dependent variable: Distance traveled on the person-day level including zeroes (panel a), restricted to positive observations (panel b) and as a dummy denoting a positive daily distance. Standard errors (in parentheses) are clustered at the participant level. Estimation with PPML (panels a-b) and linear probability (panel c). The results show proportional effects (with 1.00 indicating no effect). All regressions include individual, calendar day and day of study FE, plus weather controls (not shown).

Table A.6: Distances by mode, with and without corrections

(a) Corrected data

Mode	Overall					Intensive				
	Observations	Mean (km)	Std. dev. (km)	Min. (km)	Max. (km)	Observations	Mean (km)	Std. dev. (km)	Min. (km)	Max. (km)
Car	167,916	34.95	48.20	0	498.57	134,715	43.57	50.20	0	498.57
Public Transport	167,916	10.16	31.87	0	499.22	65,360	26.11	46.83	0	499.22
Bicycle	167,916	0.62	3.54	0	127.81	14,720	7.02	9.90	0	127.81
Walking	167,916	1.71	2.25	0	20.00	138,146	2.07	2.32	0	20.00
Total distance	167,916	47.44	55.37	0	618.00	160,581	49.60	55.67	0	618.00

(b) Uncorrected data

Mode	Overall					Intensive				
	Observations	Mean (km)	Std. dev. (km)	Min. (km)	Max. (km)	Observations	Mean (km)	Std. dev. (km)	Min. (km)	Max. (km)
Car	168,362	34.96	48.67	0	498.57	131,082	44.90	50.95	0	498.57
Public Transport	168,362	10.05	32.44	0	499.22	56,652	29.88	50.34	0	499.22
Bicycle	168,362	0.62	3.61	0	128.57	14,058	7.40	10.31	0	128.57
Walking	168,362	1.72	2.27	0	20.00	138,573	2.09	2.34	0	20.00
Total distance	168,362	47.35	55.40	0	618.00	160,974	49.52	55.70	0	618.00

Notes: The table shows the summary statistics by mode on the person-day level. The "uncorrected" data correspond to the data as imputed by the app, without considering any corrections made by the users. The number of observations differs slightly due to cleaning steps that remove implausible data based on average speed (see Section 3.1 in the main text).

Table A.7: ATE on imputed travel distance (no corrections)

(a) Overall margin

	Total	Car	PT	Bicycle	Walking	Slow
Pricing	0.982	0.967'	1.014	1.032	1.038*	1.043
	(0.014)	(0.017)	(0.039)	(0.084)	(0.019)	(0.027)
Information	1.006	0.983	1.086*	0.976	1.008	1.008
	(0.014)	(0.016)	(0.042)	(0.065)	(0.018)	(0.022)
Difference	0.975'	0.983	0.934*	1.058	1.031'	1.035
	(0.014)	(0.017)	(0.033)	(0.089)	(0.018)	(0.027)
Adj. R^2	0.270	0.294	0.397	0.462	0.267	0.341
Clusters	3,616	3,616	3,584	2,437	3,616	3,616
N	167,915	167,915	166,580	166,580	167,915	167,915

(b) Intensive margin

	Total	Car	PT	Bicycle	Walking	Slow
Pricing	0.983	0.976	1.017	0.966	1.040*	1.042'
	(0.014)	(0.016)	(0.035)	(0.051)	(0.017)	(0.025)
Information	1.007	0.992	1.076*	0.989	1.010	1.015
	(0.014)	(0.015)	(0.038)	(0.045)	(0.017)	(0.021)
Difference	0.976'	0.984	0.946'	0.976	1.029'	1.026
	(0.014)	(0.016)	(0.031)	(0.055)	(0.017)	(0.025)
Adj. R^2	0.277	0.296	0.450	0.623	0.261	0.358
Clusters	3,616	3,614	3,536	1,763	3,615	3,615
N	160,580	134,712	65,310	14,041	138,145	139,981

(c) Extensive margin

	Total	Car	PT	Bicycle	Walking	Slow
Pricing	0.999	0.989'	0.992	1.028	1.005	1.008
	(0.003)	(0.006)	(0.016)	(0.044)	(0.006)	(0.006)
Information	0.998	0.989'	1.001	0.978	1.002	1.003
	(0.003)	(0.006)	(0.015)	(0.041)	(0.006)	(0.006)
Difference	1.001	1.000	0.990	1.051	1.003	1.005
	(0.003)	(0.007)	(0.015)	(0.046)	(0.006)	(0.006)
Adj. R^2	0.002	0.023	0.107	0.192	0.017	0.016
Clusters	3,616	3,616	3,584	2,437	3,616	3,616
N	167915	167,915	166,580	114,209	167,915	167,915

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. These are the same regressions as in table A.5, but using the uncorrected data. For additional notes, see table A.5.

Table A.8: ATE on travel distance: Linear model

(a) Overall margin

	Total	Car	PT	Bicycle	Walking
Pricing	-1,013.27	-1,618.08**	483.48	61.04	60.30*
	(695.21)	(598.72)	(408.07)	(53.08)	(30.53)
Information	183.13	-988.99’	1,130.58**	31.46	10.08
	(669.80)	(576.84)	(383.28)	(45.12)	(30.51)
Distance	49,998	45,263	30,186	7,242	2,041
Proportional effect					
Pricing	-0.020	-0.036**	0.016	0.008	0.030*
Information	0.004	-0.022’	0.037**	0.004	0.005
N	168,362	168,362	168,362	168,362	168,362

(b) Intensive margin

	Total	Car	PT	Bicycle	Walking
Pricing	-995.91	-1,298.60’	461.16	-231.50	78.85*
	(710.43)	(699.46)	(1,092.77)	(429.10)	(34.59)
Information	207.40	-695.60	2,519.63*	101.13	21.08
	(686.05)	(681.28)	(1,055.09)	(337.36)	(34.25)
Distance	47,967	35,430	10,293	510	1,689
Proportional effect					
Pricing	-0.021	-0.037’	0.045	-0.454	0.047*
Information	0.004	-0.020	0.245*	0.198	0.012
N	160,974	131,081	56,493	13,432	138,572

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. To derive the proportional effect as in Table A.5, we divided the absolute effects (in meters/day) by the average distance of the control group during the treatment phase and add 1. All regressions include individual, calendar day and day of study FE plus weather controls (not shown).

Table A.9: ATE on departure time for car trips

	Any time of day	Only morning	Only evening
Pricing	1.259	-4.424*	1.598
	(2.559)	(1.969)	(2.184)
Information	-2.802	-2.149	-0.123
	(2.510)	(2.011)	(2.213)
Difference	4.061	-2.275	1.721
	(2.547)	(2.101)	(2.226)
Adj. R^2	0.054	0.210	0.119
Clusters	2,962	2,955	2,960
N	280,042	100,685	179,350

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

(a) Mode Distance Share

	Car	Public transport	Bicycle	Walking
Pricing	0.970**	1.041’	1.173*	1.057*
	(0.008)	(0.023)	(0.077)	(0.028)
Information	0.985*	1.040’	1.068	1.011
	(0.007)	(0.022)	(0.066)	(0.026)
Difference	0.985*	1.001	1.098	1.045’
	(0.008)	(0.021)	(0.074)	(0.027)
Adj. R^2	0.054	0.203	0.304	0.124
Clusters	3,615	3,496	2,210	3,616
N	160,962	155,917	99,679	160,974

(b) Congestion and crowding per km

	Congestion (car)	Crowding (PT)
Pricing	0.936**	0.963
	(0.020)	(0.039)
Information	0.964’	0.935
	(0.021)	(0.039)
Difference	0.971	1.030
	(0.020)	(0.041)
Adj. R^2	0.028	0.038
Clusters	3,614	2,48
N	131,081	48,758

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

Table A.11: Mode choice mixed logit model

	Mixed logit
Alternative specific constants (base: car)	
ASC PT	0.280***
ASC bike	0.158’
ASC walk	0.542***
Random cost and travel time parameters	
Total cost (mean)	-3.039***
Total cost (sd)	0.002***
Car travel time (mean)	-4.104***
Car travel time (sd)	1.580***
PT travel time (mean)	-0.645***
PT travel time (sd)	0.002***
Bike travel time (mean)	-0.651***
Bike travel time (sd)	0.018***
Walk travel time (mean)	0.783***
Walk travel time (sd)	0.770***
Weather	Yes
Trip purpose	Yes
First mode	Yes
Departure time	Yes
N trips	171,800
N individuals	960
Log-likelihood	-161,608
AIC	323,280
BIC	323,601.8

Notes: ***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$, ’: $p < 0.1$. The model is estimated preference space with 1,500 draws and a negative lognormal distribution assumed for the random parameters.

Table A.12: Mean reversion

Sample	Control group	High external costs	Medium ext. costs	Low ext. costs.
Post	0.159			
	(0.104)			
Post × highext	-0.898**			
	(0.144)			
Post × lowext	0.543**			
	(0.084)			
Pricing		-0.282	-0.162'	-0.031
		(0.178)	(0.085)	(0.089)
Information		-0.059	-0.012	-0.179*
		(0.180)	(0.083)	(0.083)
Adj. R^2	0.234	0.141	0.046	0.050
Clusters	1,220	891	1,806	919
N	56,603	42,040	84,257	42,057

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. The first column regresses daily external costs of the control group on a treatment period dummy ("post") and interactions with dummies denoting the top and bottom quartiles of external costs during the observation period. Columns 2-4 use the base regression, but restricted to subsamples according to the cost quartiles.

Table A.13: Personal values by treatment group

	Altruistic	Biospheric	Egoistic	Hedonic
Pricing	-0.016	-0.059*	0.009	-0.016
	(0.026)	(0.028)	(0.027)	(0.029)
Information	0.017	0.027	0.025	-0.003
	(0.026)	(0.028)	(0.027)	(0.029)
N	3375	3375	3375	3375

Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$.

Table A.14: Observability

Dependent Variable:	Observed_{post}	Observed_{post}	External costs
$\text{External costs}_{pre}$	0.109**	0.033	
	(0.036)	(0.033)	
Pricing	0.041	0.065	-0.069
	(0.247)	(0.228)	(0.110)
Information	0.103	-0.042	-0.033
	(0.245)	(0.226)	(0.101)
Male	-0.175	-0.160	
	(0.208)	(0.192)	
Age>54	-0.289	-0.209	
	(0.264)	(0.244)	
Age<30	-0.000	-0.264	
	(0.262)	(0.243)	
French sp.	-0.257	-0.263	
	(0.232)	(0.215)	
Suburban	0.136	-0.019	
	(0.230)	(0.213)	
Rural	0.472	0.302	
	(0.382)	(0.354)	
Car owner	-0.335	-0.221	
	(0.341)	(0.315)	
Observed_{pre}		0.628**	
		(0.025)	
Active × Pricing			-0.198
			(0.116)
Active × Info			-0.081
			(0.105)
Adj. R^2	0.002	0.147	0.234
Clusters	3,616		
N	3,690	3,690	168,362

Notes: **: $p < 0.01$, *: $p < 0.05$. In the first two columns, we regress the number of valid observations during the treatment period (ranging from 0 to 27). Observed_{pre} is the number of observations during the pre-treatment period. "Active" is a dummy denoting the 50 % of participants that recorded more than 24 observations during the pre-treatment period.

B The MOBIS study: Study design details

The study protocol is presented in detail in Molloy et al. (2023), including the invitation protocol and survey methods. Here in the appendix we include additional information on the calculation of the necessary sample size, compensation and participant support.

B.1 Determining the 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 et al. (2007), are biased (Burlig et al., 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.[39] 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 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

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

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.1 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.1: Virtual budgets, remaining balances and incurred costs (CHF).

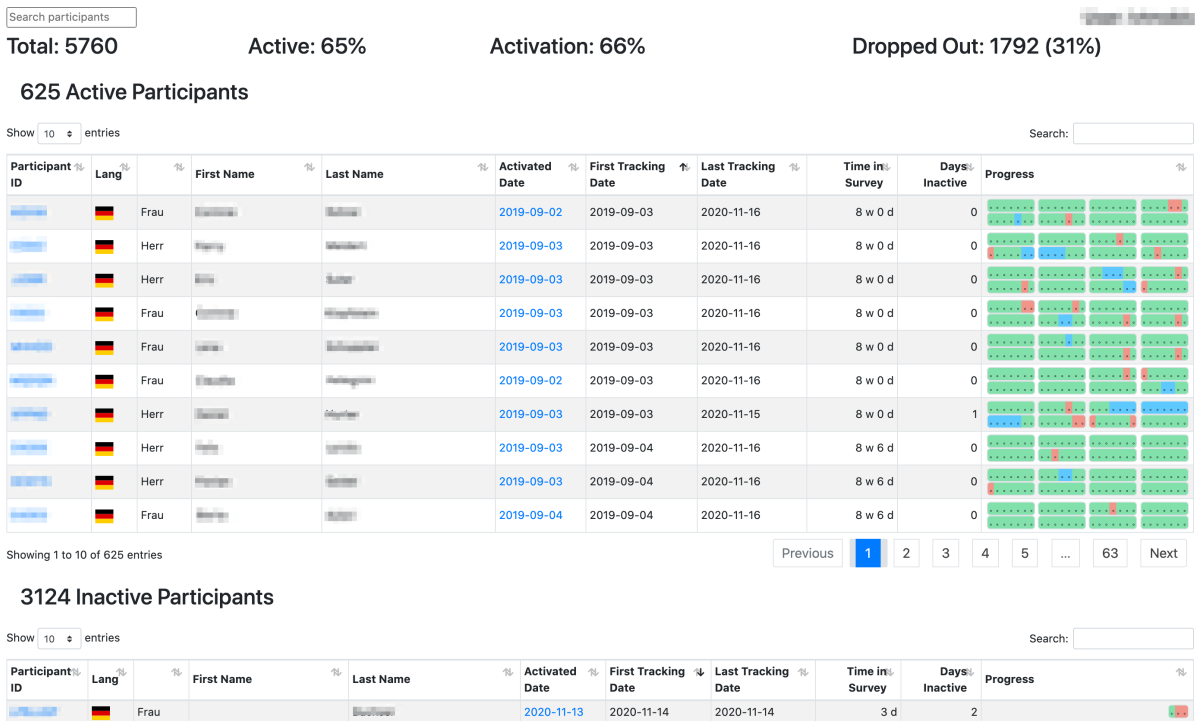
	Virtual budget	Remaining balance	Incurred costs
Mean	173.82	45.45	132.89
Std. dev.	101.63	48.53	81.66
Min	50.00	0.00	0.00
25%	100.00	7.00	75.72
50%	150.00	31.44	115.37
75%	230.00	68.53	172.72
Max	745.00	432.68	616.08

B.3 Study monitoring and user support

Two dashboards were developed for the monitoring of both the participants and the participation rate (see Figures B.1 and B.2 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. Figure B.3 shows the number of participants starting with the tracking, by calendar week.

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.

Figure B.1: Overview page of participants



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

Figure B.2: Screenshot of the MOBIS response rates dashboard

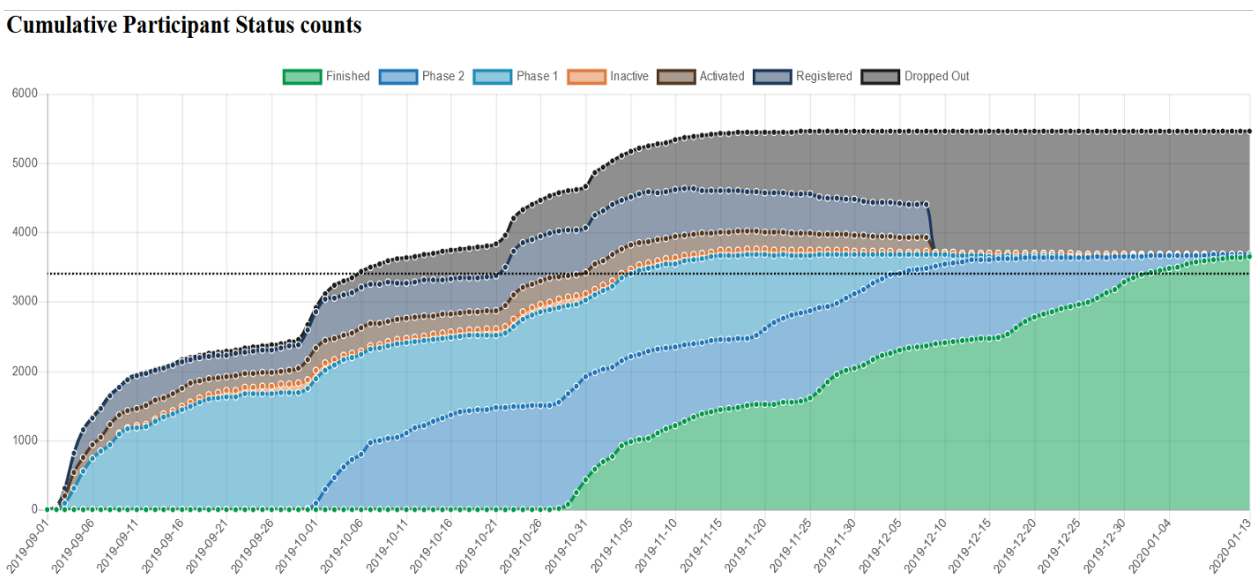
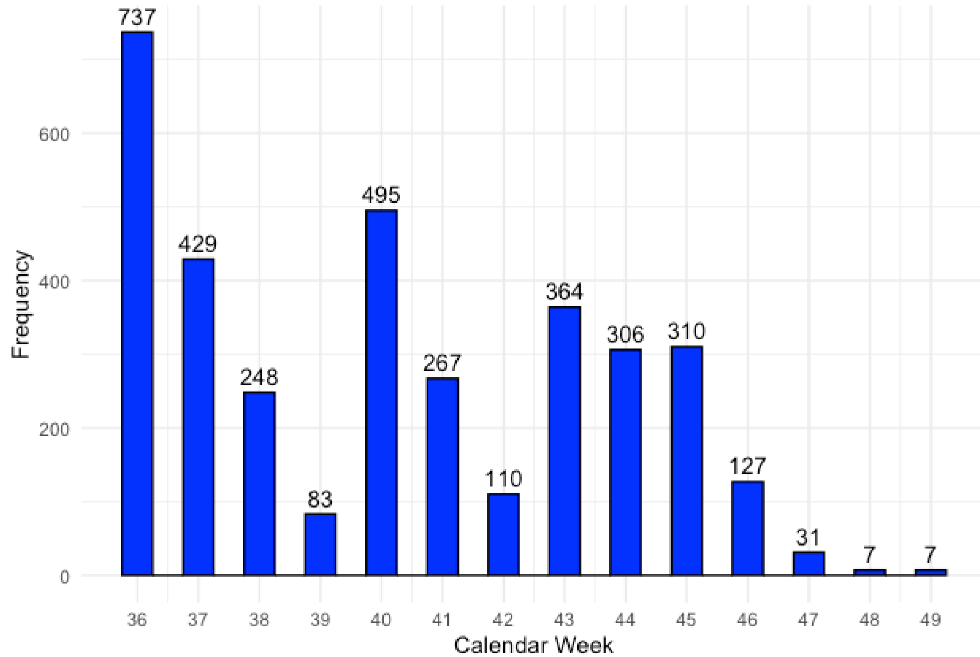


Figure B.3: Starting participants, by week



Notes: This figure shows the number of people that started tracking, by calendar week of 2019.

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.

References

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

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, May.

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

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

Molloy, Joseph, Alberto Castro, Thomas Götschi, Beaumont Schoeman, Christopher Tchervenkov, Uros Tomic, Beat Hintermann and Kay W Axhausen (2023). "The MOBIS dataset: a large GPS dataset of mobility behaviour in Switzerland." *Transportation* 50(5): 1983–2007.