Pigovian Transport Pricing in Practice*

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

May 18, 2026

Abstract

We implement Pigovian transport pricing in a field experiment in urban agglomerations of Switzerland over the course of 8 weeks. Our pricing considers the external costs from climate damages, health outcomes from pollution, accidents and physical activity, and congestion. It varies across time, space and mode of transport and is deducted from a budget provided to GPS-tracked participants. The treatment significantly reduces the external costs of transport during the course of the experiment. The main underlying mechanism is a shift away from driving towards other modes, such as public transport, walking and cycling. Providing information about the external costs of transport alone is insufficient to change the transport behavior for the sample majority. A time-invariant tax on CO_2 and health-related externalities would capture most of the welfare gains associated with the first-best policy.

*This research has been supported by the Innosuisse "SCCER CREST/Mobility Joint Activity" and the Swiss Federal Roads office (grant no. 0820001327). We thank A. Bento, M. Fowlie, E. Glaeser, B. Harstad, R. Kellogg, G. Kreindler, J. List, T. Rutherford, K. Schmidheiny, A. Stutzer, C. Wunsch and seminar participants at ETH Zurich, Geneva Graduate Institute, Georgetown University, Harvard University, HEC Lausanne, NBER, RWI Essen, TU Graz, UC Berkeley, UC Davis, UC Santa Barbara, University of Basel, University of Chicago, University of Neuchatel, UW Madison, WU Wien, participants at various conferences and workshops, and four anonymous reviewers for valuable comments. Research assistance was provided by C. García and J. Roth, and proofreading by M. Peters. The usual disclaimer applies.

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

‡Univ. of Hamburg, Faculty of Business, Economics & Social Sciences, Von-Melle-Park 5, 20146 Hamburg, Germany.

§FAIRTIQ AG, Aarbergergasse 29, 3011 Bern, Switzerland.

¶Univesity of Oregon, School of Planning, Public Policy and Management, 97403 Eugene, OR, USA.

‖Swiss Tropical and Public Health Institute, Kreuzstrasse 2, 4123 Allschwil, Switzerland.

**Ville de Lausanne, Division mobilité, Rue du Port-Franc 18, 1001 Lausanne, Switzerland.

††ZHAW, Institute of Sustainable Development, Technoparkstr. 2, 8400 Winterthur, Switzerland.

‡‡ETH Zürich, IVT, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland.

Keywords: Transport pricing; Pigovian taxation; mobility; transportation; external costs; congestion; GPS-tracking.
JEL Codes: H23, H31, I18, Q52, Q54, R41, R48

1 Introduction

Transport systems face multiple challenges. In many cities around the world, drivers lose over 100 hours per year due to traffic congestion (INRIX, 2020). Public transport can help reduce congestion (Anderson, 2014), but increasing the capacity of private and public transport faces physical limitations and high costs due to competition with other land use. 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 marginal external costs of transport.[1] In this paper, we implement a multi-modal transport pricing scheme, based on the values suggested by the Swiss government, and estimate its effects on individual transport choices. The main modes considered are car, public transport, cycling and walking, and their combined external costs are around CHF 4.50 per person and day.[2]

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 around 3,300 people living in urban agglomerations of the German- and French-speaking regions of Switzerland who regularly drive and also have access to public transportation. 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 4.6%. 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. A mediation analysis indicates that a reduction in driving is responsible for close to 80% of the total effect.

To differentiate the pricing effect from a possible effect due to the information that is embedded in the pricing, 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. This "pure information" effect is not statistically significant or the overall sample. However, we do observe a significant reduction in the external costs from participants with an above-average score for an "altruistic" measure that we compute using a questionnaire. The differential effect between the pricing and information groups can be interpreted as the causal effect of adding a financial

[1] For a discussion of the definition of the external costs of transport, see CE Delft (2019).

[2] At the start of the study (September 2019), one Swiss franc (CHF) corresponded to 0.92 euro and 1.01 US dollars.

incentive for people who already receive information 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 a role in explaining the total response to the intervention, but that the pricing effect dominates.

We approximate the welfare effects from our transport pricing by estimating a discrete-choice model on the trip level, thus focusing on mode choice. 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 24 per person and year. Due to additional substitution possibilities, the long-run effect is likely larger. Using our model parameters, we can construct counterfactual policies that raise the same revenue. We estimate that a CO2 tax would capture 35% of the total benefits, and that adding a volumetric tax on health-related externalities would achieve 86% of the benefits of the first-best policy. The welfare estimates are sensitive to the main underlying cost parameters, such as the social cost of carbon, the value of a statistical life or the value of time. The cost parameters used in this study are taken from the Swiss Federal Office of Spatial Development.

Our discrete choice model can also be used to derive the value of travel time and price elasticities by mode. We find that people are willing to pay CHF 10 to save one hour of driving, whereas the corresponding values for public transport, bicycling and walking are CHF 37, 32 and 43, respectively. Driving is associated with an own-price elasticity of -0.21 in terms of distances, whereas the corresponding estimate for public transport is -0.25.

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) and Vickrey (1963). In first-best pricing, the full marginal external cost is charged to all users, who will then internalize it 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).

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).

A number of experimental studies has exposed study participants to individualized pricing schemes (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.

Only few RCTs of a similar scale to ours 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 (2024) measures 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 (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. 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, and to information, in a multi-modal context. Third, by estimating a structural model of transport demand, we can approximate the welfare effects of our intervention, as well as of other price-based interventions that have a lower information requirement.

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.

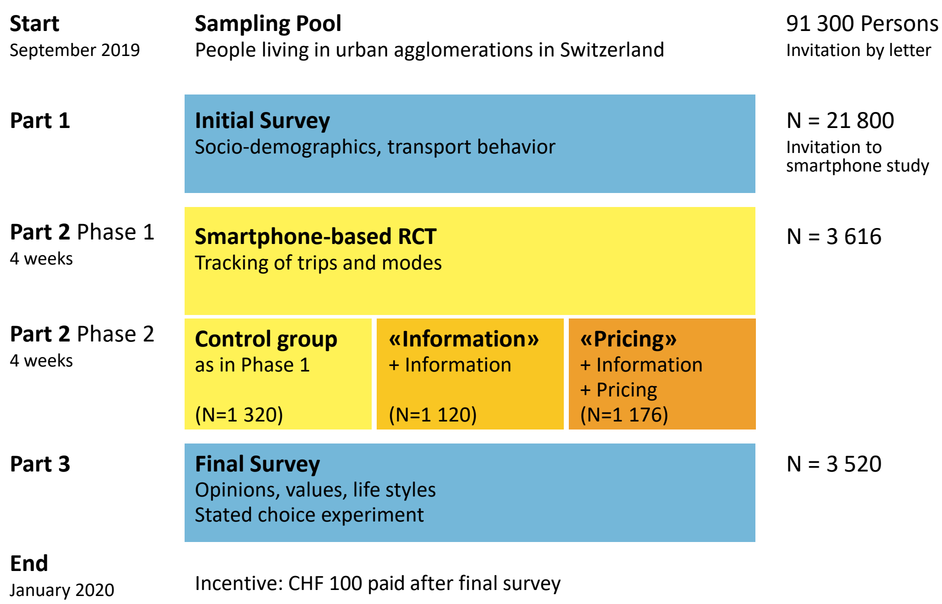
2 The MOBIS experiment

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 C.

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.[3]

Figure 1: Design of the MOBIS experiment



The first part of the study 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. Two-thirds of those actually started tracking. The third part consisted in a final 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

[3]We were provided with 60,000 addresses from the Federal Office of Statistics at no charge. When it became clear that this would not be sufficient to recruit the 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. 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. Randomization worked well, with most variable being balanced across the three groups in the RCT sample (see Fig. B.1). 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. Table 1 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 and geographic area as our study. The respondents of the introduction survey are very similar to the MTMC population. The largest differences are in terms of the share of young adults, education and nationality.[6] 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 share of people that have access to a car is higher in the RCT sample than in the MTCM as conditioned participation on driving regularly. We discuss the implications of our sample selection procedure for external validity in section 6.

[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).

Table 1: Demographic information for the MOBIS sample

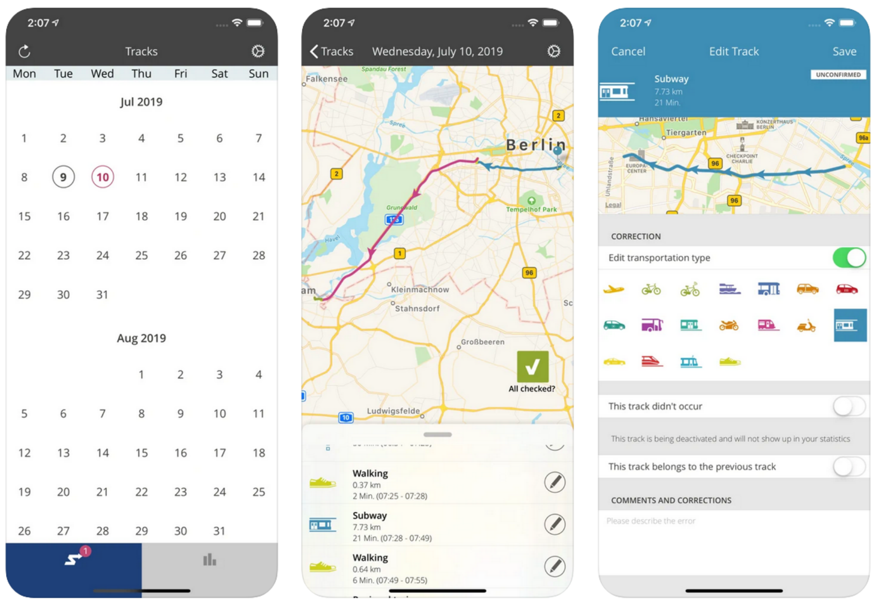
Variable	Level	MOBIS Intro	MOBIS Tracking			MTMC
			Control	Info	Pricing	
Age	[18, 25]	20.1	18.2	19.8	19.7	14.3
	(25, 35]	19.4	17.9	18.6	16.8	21.4
	(35, 45]	19.9	22.2	21.0	24.4	22.6
	(45, 55]	21.6	23.1	24.2	22.7	23.7
	(55, 65]	19.0	18.6	16.4	16.3	17.9
Education	Mandatory	9.2	8.0	5.1	6.8	13.8
	Secondary	43.3	47.3	49.3	48.3	47.5
	Higher	47.5	44.7	45.5	44.9	38.7
Employment	Employed	68.7	73.5	71.9	70.9	68.8
	Self-employed	7.3	6.3	5.2	7.2	8.8
	Apprentice	1.9	1.9	1.6	1.6	2.2
	Unemployed	4.4	3.3	3.9	4.7	3.9
	Student	9.3	7.6	8.7	7.9	3.0
	Retired	2.5	2.6	2.2	2.3	3.6
	Other	5.9	4.7	6.5	5.4	9.7
Gender	Male	48.9	50.3	50.0	49.3	49.4
	Female	51.1	49.7	50.0	50.7	50.6
Household size	1	15.5	11.3	11.4	12.0	18.3
	2	31.7	30.1	30.9	28.5	32.0
	3	20.5	22.8	21.5	19.9	19.9
	4	23.6	25.3	28.0	30.1	20.7
	5 or more	8.6	10.5	8.2	9.4	9.1
Income	4,000 CHF or less	12.2	6.6	8.3	7.3	8.8
	4,001 - 8,000 CHF	29.4	30.8	29.7	27.0	31.4
	8,001 - 12,000 CHF	24.5	28.1	30.1	30.2	24.6
	12,001 - 16,000 CHF	12.1	15.7	13.7	14.5	11.7
	More than 16,000 CHF	8.0	9.6	9.4	10.7	8.4
	Prefer not to say	13.8	9.1	8.7	10.3	5.8
	Don't know					9.2
Language	German	62.7	66.3	65.3	66.4	69.5
	French	28.6	26.1	26.5	26.4	26.5
	Italian					4.0
	English	8.7	7.6	8.2	7.2	
Nationality	Switzerland	78.1	81.4	80.6	82.3	69.5
	Other	21.9	18.6	19.4	17.7	30.5
Area	Urban	75.0	63.5	64.3	63.9	77.4
	Intermediate	18.1	28.2	27.3	28.3	16.6
	Rural	6.8	8.3	8.4	7.7	6.0
Access to car	Yes	61.0	87.1	88.0	87.9	69.7
	Sometimes	15.5	11.8	10.5	11.1	22.7
	No	23.5	1.2	1.5	0.9	7.5
Full PT subscription	Yes	37.2	21.9	25.1	25.3	34.5
Half fare PT subscription	Yes	47.6	49.0	49.1	48.1	37.6
No PT subscription	Yes	26.0	33.8	32.5	33.9	37.9
Access to bicycle	Yes	68.5	72.7	72.1	69.6	70.1
	Sometimes	4.1	4.4	5.5	3.9	8.8
	No	27.4	22.9	22.4	26.5	21.1
N		20,783	1,205	1,208	1,158	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. See Figure B.1 for information on covariate balance between the treatment groups based on standardized mean differences.

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. 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, and 5% of the modes were corrected. The database stores both their correction and the original algorithmic imputation. This algorithm achieved an overall accuracy of over 90% (see Molloy et al., 2023, Table 8).

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

[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, 2017). Hence, we made validation of the trip purpose and mode optional for participants, in order to not increase the response burden excessively over the 8 weeks.

car, ferry, train (local, regional and long-distance), tram and walk. In addition, users could select the following modes as a correction: Boat, car sharing, gondola, motorbike/scooter, taxi/Uber. 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. Since the app cannot differentiate between drivers and passengers, participants could not avoid costs by car-pooling instead of driving themselves.

2.3 The external costs of transport

The assessment of the external costs of transport in this paper 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. and will therefore vary across jurisdictions and time, depending on the policy mix in place and the valuation of the external cost components. In Switzerland, the costs of road maintenance are paid for by an excise tax on fuels and are therefore already internalized.[9]

The non-internalized costs are shown in Table 2 in cents per km. They can be grouped into climate damages, congestion, and health-related costs. The latter includes health damages associated with local pollution and noise as well as accident-related health costs. These are external to the involved persons due to the socialized accident insurance system in Switzerland. The health dimension also includes health care savings due to improved health from active transport (Götschi et al., 2016). Throughout the paper, we focus on the *marginal external* costs associated with an individual trip and exclude the life-cycle emissions associated with producing cars and building road and rail infrastructure, as well as all internal costs and benefits.

We computed the external costs for the recorded daily trips using an automated data pipeline that included also data collected from the online introduction survey (e.g., engine type and size). For the calculation of external costs associated with driving, a partial-equilibrium approach described in detail 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

[9]The Swiss fuel tax (currently CHF 0.77 per liter of gasoline and CHF 0.80 per liter of diesel) is earmarked for road construction and maintenance. According to ARE's calculations, private motorized transport generated a total cost of about CHF 52.5 billion in 2019. Road users contributed a total of CHF 45 billion (in the form of fuel and other taxes), whereas the remaining CHF 7.5 billion were not internalized but borne by society at large in the form of emissions and health costs (Federal Statistical Office, 2022, Fig. G4). These are the external costs that we focus on in our study.

Table 2: Average external costs by mode (in Swiss cents per person-km)

Mode	CO2	Pollution & noise	Accident health costs	Health benefits	Con- gestion	Crow- ding	Total
Car	2.43	5.36	2.01	-	2.44	-	12.23
Bus	1.44	3.77	1.41	-	-	1.18	7.81
Intercity train	0.01	0.91	0.07	-	-	1.01	1.99
Regional train	0.01	0.97	0.07	-	-	1.82	2.86
Tram & lightrail	<0.01	0.16	1.26	-	-	1.37	2.79
Bicycle	-	-	25.70	-18.71	-	-	6.99
Walk	-	-	7.50	-18.62	-	-	-11.12

Notes: The values for public and active transport are based on NISTRA (Federal Roads Office, 2017). Crowding costs for public transport of CHF 0.1/km were applied for congested links (see text). Negative costs (walking) indicate an external benefit. The external costs of driving vary over time and space and were computed within MATSim (Molloy et al., 2021).

to calculate the external costs of congestion and emissions. The emissions factors were taken from the HBEFA database 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, CHF 515 (1,358) per kg of PM_{10} in rural (urban) areas and 7,109 CHF per ton of NO_x and a value of travel time of CHF 26 per hour. 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 urban densities.

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). As a practical solution, a peak-hour pricing scheme 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. We then applied a peak surcharge of 0.10 CHF/km to PT trips between any two municipalities (or within a single municipality) for which peak demand exceeds offpeak

[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.

demand by more than a factor of 3.[11] Throughout the experiment, participants had access to an interactive map which showed them where and when the PT crowding scheme applied.

For modes other than driving, the per-km values presented in Table 2 were directly applied to the recorded length of the trip, including the crowding fee if applicable. Whereas walking is associated with net external benefits, the external accident costs outweigh the external health benefits from cycling.[12]

In principle, we could have used any pricing scheme and estimated the participants' response to it. We chose the Pigovian rate (or, at any rate, an estimate of it) for two reasons.[13] 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 facilitates the welfare calculations below. Besides this benchmark, we also provide estimates of the welfare implications if the actual pricing were to deviate from the Pigovian rate.

Last, we stress that these external cost estimates are average calculations that are subject to a number of assumptions, and they do not always incorporate all relevant heterogeneity. For example, whereas the pollution and climate cost estimates depend on the car size and type, we use the same accident externalities for all cars (despite the fact that larger cars cause more severe accidents than smaller ones; see Anderson and Auffhammer, 2014) and irrespective of the weather or the socio-demographic characteristics of the driver. Similarly, our congestion externality is measured with error and may thus assign a congestion cost for a particular trip that is absent in reality, and vice versa. For these reasons, our pricing is an approximation of "first-best" on average, but it cannot capture the true external cost of every trip.

[11] The peak hour windows and the affected municipality-pairs were determined using the MATSim scenario output for Switzerland (Bösch et al., 2016). If a 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.

[12] Most of the positive health effects from cycling are private in the form of lower morbidity and mortality and at least partly internalized by cyclists (Götschi and Hintermann, 2014). In our pricing, we only include external costs and benefits as they arise via the health insurance system in Switzerland.

[13] 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.

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.[14] 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). Even for people with pre-existing knowledge about the external costs of transport, it would have been difficult to know the exact magnitude of their own costs based on the ARE methodology. In this sense, the information contained some novel aspects for everyone.

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.[15] This treatment thus simulated transport pricing based on the monetized marginal external costs of transport.

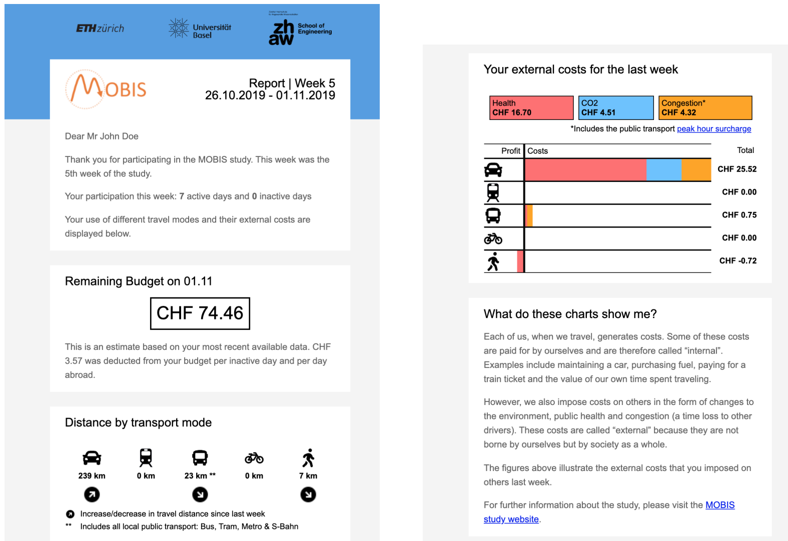
The nested design of the treatments allows for an estimation of the effect of "pure money" in the sense of providing a monetary incentive to individuals that are perfectly informed about their external costs of transport.[16]

[14] 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. 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.

[15] 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. To remove a financial incentive to disable the app, the participants were informed that their average pre-treatment costs would be subtracted from the budget for each missing day.

[16] 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 treatment 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, 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 B.1 provides a proportional breakdown of the external costs by 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. Most external costs of transport are local, with climate 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.[17]

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 (i) by construction, and since our RCT sample consists of around 3,600 individuals out of

[17]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

						Control vs. Info.	Control vs. Pricing
Dimension	Outcome	Unit	Control	Information	Pricing	p	p
External costs	Total	CHF/d	4.50	4.58	4.70	0.409	0.083
			(5.68)	(5.63)	(5.79)		
	Congestion & crowding	CHF/d	1.03	1.07	1.14	0.243	0.004
			(1.59)	(1.58)	(1.69)		
	Climate	CHF/d	0.88	0.88	0.90	0.854	0.503
			(1.29)	(1.28)	(1.29)		
	Pollution & noise	CHF/d	1.97	2.01	2.03	0.442	0.220
			(2.59)	(2.57)	(2.62)		
	Health care costs	CHF/d	0.61	0.62	0.63	0.580	0.371
			(1.02)	(1.02)	(1.05)		
Private costs	Total	CHF/d	22.30	22.71	23.17	0.439	0.231
			(31.37)	(30.95)	(32.37)		
Tracking	Distance	km/d	46.78	47.88	49.31	0.317	0.025
			(55.26)	(54.62)	(57.26)		
	Duration	min/d	92.44	93.22	94.02	0.482	0.195
			(84.58)	(80.03)	(82.78)		
	Tracking days	Nr.	23.64	23.91	23.68	0.065	0.801
			(3.74)	(3.44)	(3.72)		
	Trips	Nr./d	4.84	4.88	4.88	0.566	0.483
			(1.60)	(1.61)	(1.48)		

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

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:

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

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 and information treatment, respectively, are active for person i on a given day, and zero otherwise. 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. 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).

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 for everyone in Switzerland, whereas the day-of-study FE account for the possibility that respondents change their behavior in response to being tracked. The combination of both time-related 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. We allow for a correlation of the error term ϵ_{its} within participants, but not between. To address the concerns involving two-way FE estimators (see, e.g., De Chaisemartin and d'Haultfoeuille, 2020), we also apply the "interaction-weighted" estimator developed by Sun and Abraham (2021). However, because the results are almost identical, we rely on our base specification throughout the paper.[18]

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 from MeteoSwiss provided on a 1 x 1 km grid.[19] 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

[18]Many of the concerns arising from observational studies with two-way FE are attenuated in our setting, as (i) the treatment is randomized, (ii) there is a never-treated control group, and (iii) all units are observed for the same time period before and after treatment. However, since it is possible that people who responded quickly to our first invitation letter have different potential outcomes as those that responded with a delay, the ATE as computed by eq. 1 may be biased. Comparing Table B.3 with our base results in Table 4 shows that this is not the case.

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

"Cold" for an observed trip j on day t in terms of the daily maximum or minimum relative to threshold values (see appendix).

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.[20]

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. The weather variables are jointly highly significant and affect external costs mainly via distance traveled (see Table B.5). However, as their inclusion does not change the ATE, we refrain from showing both versions for the remainder of the paper.

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. 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 (denoted as "Difference") is statistically significant only for congestion. Furthermore, the effect is immediate and does not significantly change over the course of the treatment period (see Tables 11 and B.3 for the effects by treatment week).

To interpret the magnitude of the ATE, we can compare the proportional response in external costs to the change in the total transport price, including both private and external

[20]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).

Table 4: Average treatment effects on external costs

	Total ext. costs		Health costs		Climate costs		Congestion costs	
Pricing	-0.229**	-0.230**	-0.119**	-0.120**	-0.039*	-0.039*	-0.071**	-0.071**
	(0.070)	(0.070)	(0.044)	(0.043)	(0.016)	(0.016)	(0.022)	(0.022)
Information	-0.094	-0.098	-0.051	-0.054	-0.021	-0.022	-0.022	-0.022
	(0.067)	(0.067)	(0.042)	(0.042)	(0.015)	(0.015)	(0.021)	(0.021)
Difference	-0.135'	-0.133'	-0.068	-0.067	-0.018	-0.017	-0.049*	-0.049*
	(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		-0.498**		-0.357**		-0.129**		-0.012
		(0.073)		(0.048)		(0.017)		(0.018)
Cold		0.188**		0.158**		0.058**		-0.027**
		(0.018)		(0.012)		(0.004)		(0.004)
Adj. R^2	0.232	0.234	0.225	0.227	0.222	0.224	0.265	0.266
Clusters	3,616	3,616	3,616	3,616	3,616	3,616	3,616	3,616
N	168,359	168,359	168,359	168,359	168,359	168,359	168,359	168,359

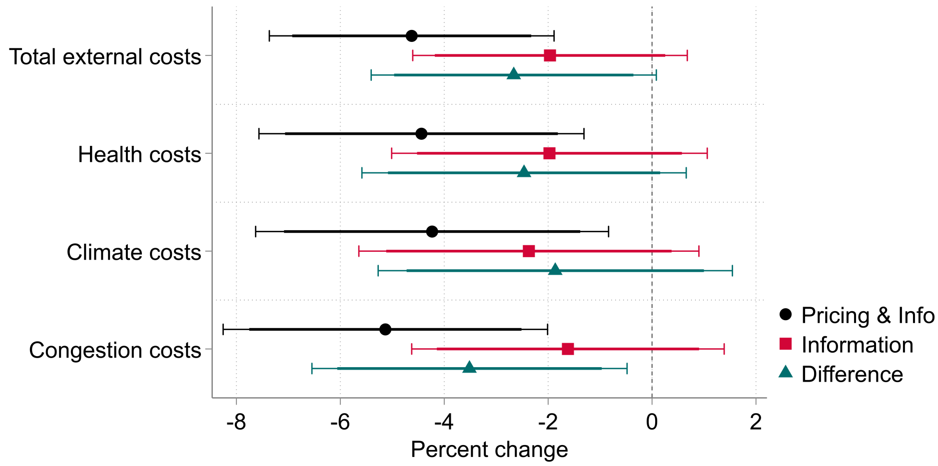
Notes: **: $p < 0.01$, *: $p < 0.05$, ': $p < 0.1$. The dependent variable is the external cost of transport aggregated to the person-day level. Standard errors (in parentheses) are clustered at the participant level. All regressions include fixed effects on the person, calendar-day and day-of-study level. Standard errors for proportional effects and elasticities are bootstrapped.

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 cost of 59 cents/km, with an inter-quartile 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 these assumptions, we obtain a transport tax-related price increase of 19.3%. The proportional reduction in the external costs is 4.6%. Dividing the latter by the former results in -0.24, which describes the short-term elasticity of external costs in response to a

[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 (which is the main reference category in Switzerland). 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. The savings implicit in the subscription ranges from 24% in Geneva to 76% in Basel.

[22]This question was part of the final survey and asked: "What is the average private cost of your car travel per kilometer?" 90% of the respondents answered this question. 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%. Finally, we imputed the missing values based on a linear model associating private costs with information about the age, size and fuel type. Note that these costs may be an underestimate of the true costs, as reported by Andor et al. (2020). However, since people make choices based on their beliefs, it is the expected rather than the actual costs that matter in this context.

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 dividing 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.

one-percent increase in the costs of transport. Note, however, that this estimate does not identify a fundamental behavioral parameter but is specific to our pricing.[23]

Table B.2 shows the sensitivity of the results with respect to the inclusion of fixed effects and the presence of a control group. Removing either the study day or the calendar day fixed effects does not significantly change the results; however, when removing both, the ATE more than doubles, suggesting that controlling for unobserved temporal shocks is important. When estimating the effects without a control group and using a before-vs-after setting instead, the ATE is significantly over-estimated because it also absorbs a part of the seasonal effects.

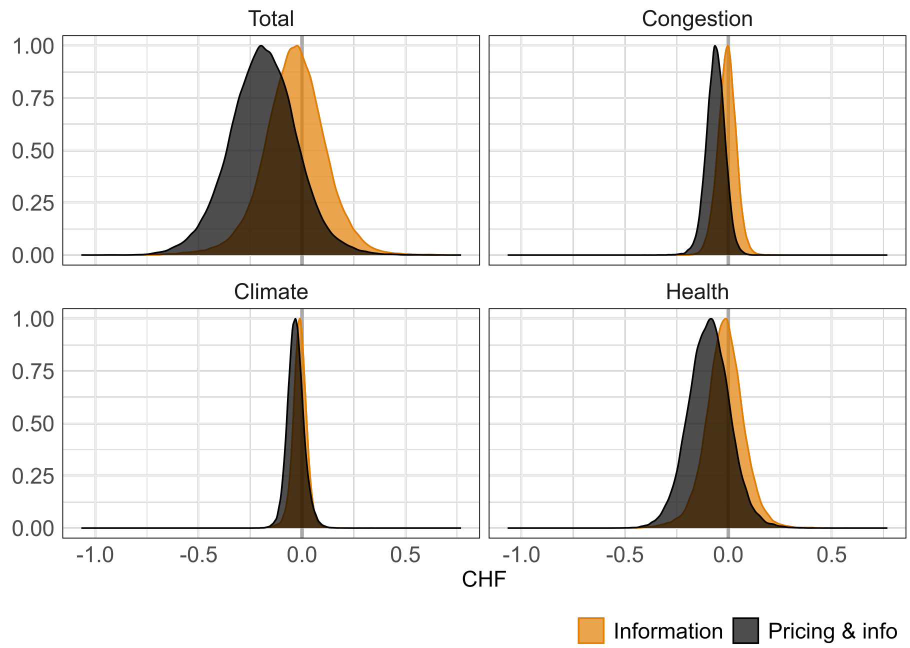
4.2 Effect heterogeneity

In order to investigate a potential effect heterogeneity, we start by employing a "causal forest" approach based on the generalized regression forest algorithm proposed by Wager and Athey (2018) and implemented by Tibshirani et al. (2020). 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

[23]For example, if the same behavioral response were to be achieved by subsidies rather than taxes, the resulting elasticity would be positive.

replacement. Figure 5 shows the distribution of the conditional treatment effect, both for the pricing and the information treatments.

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.

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. To generate a benchmark of importance, we included a continuous and a discrete random variable. The variables with a higher importance ranking than these are valid candidates to explain the effect heterogeneity, since they contain "better than random noise" information. These are shown in Figure B.2 in the appendix.

Besides the socio-demographic variables, we also included a number of variables from the final survey. A set of 16 questions was used to elicit respondents' personal values (Schwartz, 1992; De Groot and Steg, 2010) and assign them an index along four dimensions labeled "altruistic", "egoistic", "hedonic" and "biospheric".[24]

[24]For the exact questions used, see Axhausen et al. (2021, p. 198). Since the personal values were elicited only after the experiment, it is possible that they are influenced by the treatment and thus do not qualify as moderators. Table B.12 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 not turn out to be statistically significant in terms of the effect heterogeneity.

We include all variables identified to be important by the CF algorithm as dummy interaction terms with the treatment indicators in eq. (1). To account for the correlation among these variables, we include them jointly in a multi-variate regression. The regression coefficients are presented in Table B.4. Overall, we find that the effect is relatively homogeneous across socio-demographic characteristics, with some exceptions. The response is stronger (i.e., more negative) for the young, those living in rural areas and car owners. The latter two results can be explained by a higher share of driving among these subgroups. For French speakers, the effect is weaker.[25] Table 5 presents the proportional effects, price increases and resulting elasticities for the sub-samples for which we found statistically significant treatment heterogeneity.

Table 5: Response to pricing treatment for different 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	-0.046	-0.074	-0.018	0.192	0.186	0.197	-0.241	-0.388	-0.095	0.001	168,359
Age (years)											
>54	-0.063	-0.130	0.003	0.171	0.160	0.182	-0.369	-0.756	0.018	0.062	33,210
30-54	-0.029	-0.063	0.005	0.193	0.185	0.201	-0.150	-0.326	0.027	0.096	91,632
<30	-0.071	-0.127	-0.015	0.206	0.193	0.219	-0.345	-0.618	-0.072	0.013	43,511
Language											
German	-0.064	-0.096	-0.031	0.185	0.178	0.192	-0.345	-0.520	-0.170	<0.001	112,149
French	0.001	-0.052	0.053	0.208	0.197	0.220	0.003	-0.247	0.254	0.980	43,924
English	-0.049	-0.154	0.057	0.215	0.189	0.241	-0.226	-0.721	0.268	0.369	12,279
Urbanity											
Urban	-0.037	-0.074	-<0.001	0.198	0.190	0.206	-0.189	-0.376	-0.001	0.048	104,488
Suburban	-0.056	-0.107	-0.005	0.184	0.174	0.193	-0.307	-0.583	-0.031	0.029	45,423
Rural	-0.111	-0.199	-0.022	0.180	0.165	0.196	-0.613	-1.101	-0.126	0.014	13,255
Car ownership											
Yes	-0.045	-0.075	-0.016	0.191	0.185	0.197	-0.236	-0.390	-0.082	0.003	147,615
No	-0.055	-0.152	0.042	0.197	0.179	0.215	-0.279	-0.766	0.207	0.261	20,743

Notes: The lower and upper bounds reflect the 95%-confidence interval, based on a bootstrap with 1,000 replications. The last columns provide the probability that the elasticity is positive and the size of the subsample.

Last, we find that the study participants that scored above the median in terms of the altruistic index responded significantly to the information treatment. This suggests that even though providing information alone is does not measurably change the transport behavior of the overall population, it does appear to influence a sub-sample that we categorize as having higher other-regarding preferences.

[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 reflect a regional effect.

4.3 Mechanisms

People can reduce their external costs of transport in different ways. They 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 corresponding regression results, see Tables B.5 - B.10).

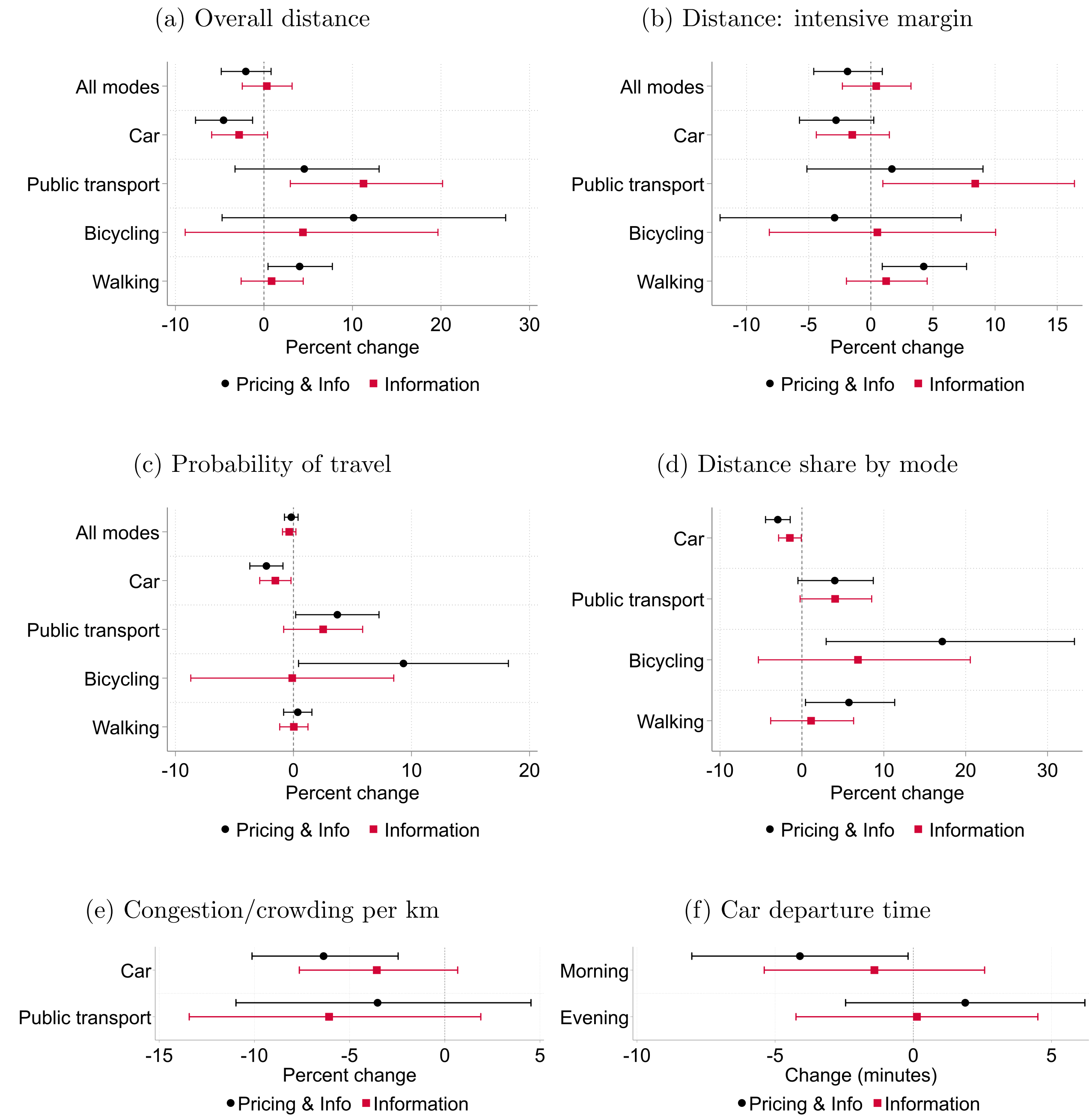
The treatment does not reduce overall travel distances for the sample as a whole, but we measure a statistically significant reduction in car distance countered by increases in the other modes. The effect can be seen separately on the intensive margin (i.e., distance conditional on traveling with a particular mode on a given day) and on the extensive margin (i.e., the probability of using a mode on a day). 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, suggesting that modal shift may not be the only mechanism responsible for the reduction in external costs. The reduction in congestion per km can be due to a selective mode shift during congested times, but also due to a change in route or departure times. Using the departure time (in minutes) as the dependent variable, we 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.[26] Note that we cannot differentiate between a shift in car departure time from a mode shift during the same time frame, as both could affect the average car departure time measured during the observation period.

To learn more about the underlying mechanisms, we engage in a mediation analysis following Baron and Kenny (1986) and Kraemer et al. (2008) (for details, see appendix). The left panel in Table 6 provides the estimates for the pricing effect that is mediated by car-km. Driving distance is indeed a powerful mediator that captures most of the treatment effect. However, there is a remaining and statistically significant "direct" effect, which is the combination of all other mediators. This effect persists despite the inclusion of an interaction term between the treatment indicator and the mediator, which allows for the relationship between car-km and external costs to differ with the treatment. For example, this would be the case if people selectively substitute car trips during highly congested times. The fact that there is a statistically significant direct effect despite this interaction term implies that the effect is not only due to mode shift, but that other mechanisms (such as a shift in routes or departure times) may also play a role. This is confirmed when using congestion costs as

[26]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).

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 car departure time (measured in minutes after midnight) as the dependent variable. The underlying regression results are shown in Tables B.5, B.9 and B.10.

the outcome variable of interest: the amount of driving does not explain the whole effect. Last, we find a positive effect associated with the information treatment mediated via public transport, suggesting that information may have contributed to the mode shift.

Table 6: Mediation analysis

	Pricing				Information			
	Coefficient	95%- percentile (lower)	(upper)	% of total	Coefficient	95%- percentile (lower)	(upper)	% of total
Direct effect on tot. ext. costs	-0.047’	-0.096	0.000	21.8	0.015	-0.030	0.060	16.8
Mediated effect: car-km	-0.168**	-0.292	-0.041	78.2	-0.107’	-0.226	0.014	116.8
Direct effect on tot. ext. costs	-0.053*	-0.099	-0.009	24.8	-0.003	-0.047	0.040	3.2
Mediated effect: car-km	-0.169**	-0.295	-0.041	78.9	-0.108’	-0.229	0.015	117.3
Mediated effect: pt-km	0.008	-0.007	0.021	3.7	0.019**	0.006	0.032	20.5
Direct effect on congestion costs	-0.041*	-0.074	-0.006	69.6	-0.010	-0.047	0.024	46.1
Mediated effect: car-km	-0.018*	-0.031	-0.004	30.4	-0.011’	-0.024	0.001	53.9

Notes: The lower and upper bounds show the 95%- percentile bootstrap confidence intervals, as recommended by Tibbe and Montoya (2022). The total effect is defined as the sum of direct and mediated effects.

5 Welfare implications

In this section, we compute the welfare effects associated with levying a Pigovian-inspired transport pricing. The welfare change can be separated into three components. First, the pricing reduces external costs; this is a welfare gain. Second, people experience a reduction in utility from the introduction of taxes. A part of this reduction is due to the transport tax payment itself; this is not lost but can be used to fund new expenditure, reduce existing taxes and/or be returned to consumers. However, to the extent that consumers respond to the pricing, the monetized utility loss has to exceed the tax payment, on aggregate.[27] This difference, which is the deadweight loss of taxation, contributes negatively to welfare. Third, transport pricing reduces the volume of driving, especially during peak hours, which will lead to increased speeds and thus to reduced travel times. Since our experiment included a negligible share of people traveling in Switzerland, we abstract from this general equilibrium effect and focus on the first two components.

To quantify the (partial-equilibrium) 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

[27]To see this, we can mentally separate the process into two steps. In the first step, people have to pay the transport taxes but are held to their original transport choices. The monetized utility loss in this step exactly equals the tax payment, by construction. In the second step, people are allowed to adjust their travel choices. But they will do so only to the extent that the resulting decrease in travel utility is smaller than their monetary savings. The fact that revenue shrinks faster than utility implies that the total tax revenue will be insufficient to compensate people for their utility loss.

compute the private and external costs associated with them; see section A.3 for more details. Our experimental setting provides us with the opportunity to combine observational data with an exogenous price change. We estimate a mixed logit model that allows the valuation of travel time to vary across individuals and modes, and the valuation of costs to vary across individuals. Because we do not allow for different departure times or routes for a given trip nor include a "no travel" option,[28] our model only captures the effect of mode choice while holding the amount, location and timing of travel constant. As we show above that mode shift (and in particular a reduction in driving) is the main mediator, this method should capture the main determinants of the change in external costs and utility.

We use the following utility function for the mixed logit models:

$$\begin{aligned} U_{ijt} &= \beta_{0j} + \beta_{cost,i} \cdot \text{Cost}_{ijt} + \beta_{time,ij} \cdot \text{Time}_{ijt} + \mathbf{w}'_{it}\delta_j + \epsilon_{ijt} && (2) \\ &= V_{ijt} + \epsilon_{ijt}, && (3) \end{aligned}$$

U_{ijt} is the utility for individual i when choosing mode j for trip t, and consists of a deterministic and a random component. We include mode- and individual-specific, i.e. *mixed*, constants β_{0ij} to capture the average utility for each mode and to control for unobserved heterogeneity between modes. We allow for individual-specific preferences for travel costs ($\beta_{cost,i}$), and for time preferences that vary by both individual and mode ($\beta_{time,ij}$). We include a vector of weather variables $\mathbf{w}_{it}$[29] and the error term ϵ_{ijt} is assumed to be drawn from an Extreme Value Type I distribution.

To estimate the monetized value of travel time (VTT) for each mode, we specify the model in willingness to pay-space. This avoids having to separately calculate the distribution of the ratio of the time parameter and the cost parameter and instead provides the mean and standard deviation of this ratio directly. We thus we re-formulate (2) as follows:

$$U_{ijt} = \beta^M_{0ij} + \beta_{cost,i} \cdot \left\{\text{Cost}_{ijt} + \beta^M_{time,ij} \cdot \text{Time}_{ijt}\right\} + \mathbf{w}'_{it}\delta_j + \epsilon_{ijt}. \qquad (4)$$

The superscript M indicates division by $\beta_{cost,i}$ and $\beta^M_{time,ij} \equiv \beta_{time,ij}/\beta_{cost,i}$ represents the individual and mode-specific VTT. This is the amount individuals are willing to pay to reduce travel time with mode j by one hour.

Estimating a mixed logit model for our full set of over 700,000 trips by maximum likeli-

[28]Including such an option is standard practice for stated preference surveys. In our setting based on actual tracking data, however, there is an infinite number of non-taken trips, and it would not be clear how to identify them in a meaningful way.

[29]Weather is included only at the day-level to avoid endogeneity due to individuals "selecting" the weather conditions for a trip. We tested the inclusion of trip-specific purposes, level of urbanity, 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. However, these variables tended to degrade the model fit overall.

hood poses a computational challenge. To obtain a more manageable number, we constrain the analysis to the pricing group and employ sample reduction techniques to provide a more computationally feasible sample size of up to 20 observations per respondent across the pre and post treatment period. Since estimating a full set of random parameters is not generally feasible with our revealed preference data, we estimate random parameters for cost, Car VTT, and the PT and Walk ASCs (see A.3). Table B.11 in the appendix presents the estimated coefficients from our preferred model (MXL 4). The cost and VTT parameters are significant and have the expected sign, and the model fit is within the typical range for revealed preference data.

The panel structure of our data allows us to calculate conditional distributions for the random parameters in the mixed logit models, in the sense that we can approximate an individual's position within the unconditional distribution (see Section A.3 for more details).

Table 7 presents the estimates of the conditional and fixed VTT for each mode, respectively, and Figure B.3 shows their distributions.[30] We find that car travel is generally less price-elastic than PT travel in terms of own-price elasticity, whereas increases in price for car travel induce more mode shift than for increases in PT travel cost.

McFadden (1977) and Small and Rosen (1981) show that the expected consumer utility, or surplus, can be calculated via the observed utility V_{ijt}:

$$E\left[CS_{it}\right] = \frac{1}{-\beta_{cost,i}} \cdot \ln\left(\sum_{j=1}^{J} \exp(V_{ijt})\right) + C_{it} \tag{5}$$

The second expression is the *logsum* for individual i and trip t. We simulate these logsums for situations with and without transport pricing and subtract one from the other. This procedure drops the unidentified constant C_{it} and yields the monetized utility loss in monetary terms due to transport pricing in our experiment:

$$\Delta E[CS_{it}] = \frac{1}{-\beta_{cost,i}} \cdot \left[\ln\left(\sum_{j\in J} \exp\left(V_{ijt}^{Pricing}\right)\right) - \ln\left(\sum_{j\in J} \exp\left(V_{ijt}^{NoPricing}\right)\right)\right] \tag{6}$$

Our expression for the average monetized utility change per person and day due to transport pricing is the sum of the change in consumer surplus over all $i \in I$ participants and $t \in T$ days, the tax revenue R (defined as the external cost charge from the model-predicted mode choices $j \in J$), and the reduction in external costs X_{ijt} that are computed

[30]Reasonable VTTs are challenging to recover with revealed preference data, however, the VTTs we estimate here are of a similar order of magnitude as the official values for Switzerland (Swiss Federal Office of Spatial Development, 2024).

Table 7: Value of travel time, choice probabilities, and elasticities

VTT (CHF/h)	Mean	Median	SD	Min	Max
Car (random)	9.80	6.34	11.93	1.77	113.40
PT (fixed)	36.78				
Bicycle (fixed)	31.60				
Walk (fixed)	43.45				
Choice probabilities	**Car**	**PT**	**Bicycle**	**Walk**	
Baseline	0.581	0.128	0.035	0.256	
No pricing	0.589	0.098	0.040	0.274	
Pricing	0.577	0.103	0.042	0.278	
Price elasticities	**Choice prob.**		**Distance**		
	Car	**PT**	**Car**	**PT**	
Price for driving	-0.11	0.34	-0.21	0.77	
Price for PT	0.03	-0.19	0.05	-0.25	

Notes: The VTT values are derived from the conditional distributions of willingness to pay to save one hour of travel time. Car has a random parameter with individual-specific heterogeneity, while PT, Bike, and Walk are specified as fixed parameters in WTP space (CHF/h). The choice probabilities are presented for the baseline (actual mode shares) and the "pricing" / "no pricing" scenarios (model-predicted shares). 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 or distance.

equivalently:

$$\Delta W = \frac{1}{I}\sum_{i\in I}\left[\frac{1}{T}\sum_{t\in T}\left(\Delta E[CS_{it}] + \sum_{j\in J} R_{ijt}^{Pricing} + \sum_{j\in J}\left(X_{ijt}^{NoPricing} - X_{ijt}^{Pricing}\right)\right)\right] \tag{7}$$

Table 8 presents the results from the preferred model. We obtain an annual welfare gain of CHF 24 per person. This is comparable to the welfare estimates computed by Almagro et al. (2024) for an optimal transport policy in Chicago.

The conceptual model that implicitly underlies this exercise is a social welfare function that is utilitarian in WTP-space. This is why only averages matter in Table 8, whereas the distribution of the gains and losses is irrelevant. But in an actual implementation of externality-based transport taxation, distributional concerns would arguably matter. The tax payment increases with income, but less than proportionally; Figure 7 shows that the tax payment as a share of income decreases (left panel) such that the tax is overall regressive. Redistributing all the tax revenue on a per capita basis removes the regressive nature of the pricing (center panel). We find that returning CHF 3.94 per day would be sufficient to make the average person in the poorest quintile indifferent in terms of transport-related utility

Table 8: Welfare effects

	Unit	Sample	Weighted
Monetized utility loss	CHF/d	-4.12	-3.95
		[-4.16, -4.09]	[-3.99, -3.91]
Revenue	CHF/d	4.06	3.88
		[4.02, 4.09]	[3.84, 3.92]
External cost reduction	CHF/d	-0.13	-0.13
		[-0.14, -0.13]	[-0.14, -0.13]
Welfare gain	CHF/d	0.07	0.07
		[0.06, 0.07]	[0.06, 0.07]
Welfare gain per year	CHF/y	24.25	24.08
		[23.10, 25.52]	[22.82, 25.50]

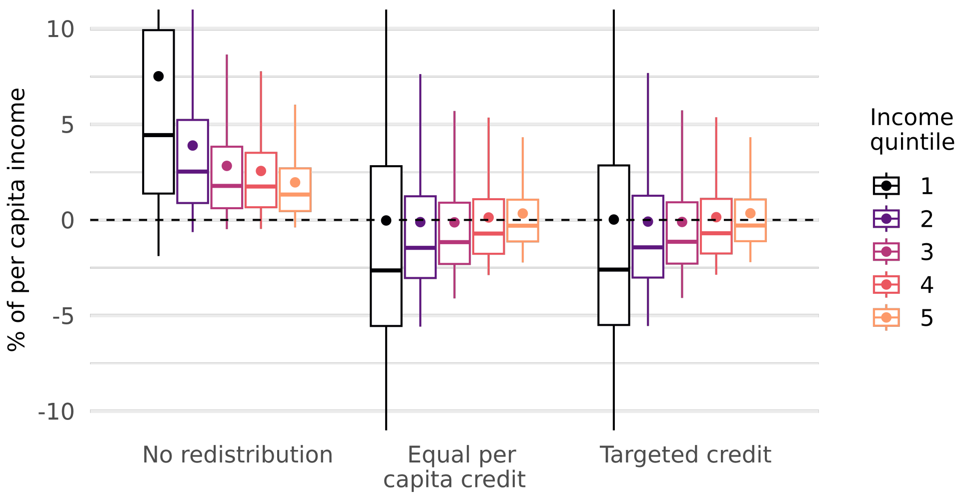
Notes: All values are averages per person. The "Sample" column shows the unweighted results for our sample, whereas the last column uses weights based on the Swiss Microcensus. The brackets show 95% confidence intervals (see A.3 for more details).

(right panel); due to the reduction in external costs, this person would be strictly better off due to the pricing. This procedure would not retain any residual revenue, however, more than 50% of individuals in all the other quintiles would be better off under this redistribution approach, which is important for political acceptance.

We can use our model to compute the welfare implications of implementing different versions of second-best pricing, which are presented in Table 9. We first implement a tax that only aims to internalize CO_2 externalities for car and public transport then supplement this with a tax on health externalities. These taxes attain 35% and 86% of the Pigovian welfare gain, respectively. We also test a perimeter pricing approach in which, in addition to the first tax, a fee is levied on car and public transport within urban areas during peak hours.[31] The perimeter tax is equal to the average congestion externality per kilometer for both car and PT trips within the perimeter. This tax achieves 42% of the Pigovian welfare gain. The fact that introducing congestion-related pricing only moderately increases the welfare gain over the first fuel tax is consistent with our observation that congestion-related costs account for a relatively small share of total transport externalities in Switzerland. And although our pricing is multi-modal and time-varying, the highest share of the external costs are due to the sheer volume of driving with internal combustion engines (see Table B.1), which can readily be addressed with a fuel tax. With a rising share of electric vehicles, of which there are very few in our sample, the fuel tax would have to be complemented with a

[31] The perimeters consist of areas defined as "urban" within medium to large agglomerations in Switzerland by the Swiss Federal Statistical Office (i.e., the two most urban settings on the nine point scale of degree of urbanity in Switzerland). The peak hours are 6:30 to 8:30 in the morning and 16:30 to 18:30 in the evening.

Figure 7: Distributive effects of Pigovian pricing



Notes: This figure presents the burden of transport pricing as a percentage by income quintile. Dot, line and box represent mean, median and inter-quartile range, resp. Left: No redistribution; middle: equal per-capita redistribution of the entire revenue; right: equal per-capita redistribution that makes the average of the poorest quintile indifferent.

per-km fee for EVs.

Throughout the paper we refer to our pricing as "Pigovian" in the sense that it is equal to the external costs computed by the Swiss federal office ARE. However, these values are subject to uncertainty, and other agencies in different countries have computed different external costs for transport. Moreover, ARE updated their own numbers after we completed our experiment, with the largest change being due to the introduction of a new methodology to compute an "equity-weighted" social cost of carbon (SCC). When using this new SCC of CHF 430/tCO_2 (Ecoplan / Infras, 2024) as opposed to CHF 136/tCO_2 in our base estimate, the welfare gain from our pricing intervention increases by 86%. More generally, we can compute the implications for welfare if the true external costs were only half, or double, of our base estimate. With the actual pricing held fixed, the welfare only changes due to adjustments in X_{ijt} in eq. (7). If the true external costs were half or double the magnitude computed by ARE and implemented in our pricing, the welfare gain would be reduced to 0% or increased to 298% of the Pigovian welfare gain.

Table 9: Welfare effects of alternative policies

	Unit	CO_2 tax	Health & CO_2 tax	Perim. & CO_2 tax	New SCC	Ext. costs halved	Ext. costs doubled
Monetized utility loss	CHF/d	-0.76	-3.05	-1.13	-3.95	-3.95	-3.95
		[-0.77, -0.75]	[-3.08, -3.02]	[-1.14, -1.12]	[-3.99, -3.91]	[-3.99, -3.91]	[-3.99, -3.91]
Revenue	CHF/d	0.76	3.01	1.13	3.88	3.88	3.88
		[0.75, 0.77]	[2.98, 3.04]	[1.12, 1.14]	[3.84, 3.92]	[3.84, 3.92]	[3.84, 3.92]
External cost reduction	CHF/d	-0.03	-0.10	-0.03	-0.19	-0.07	-0.26
		[-0.03, -0.03]	[-0.10, -0.09]	[-0.04, -0.03]	[-0.20, -0.18]	[-0.07, -0.06]	[-0.28, -0.25]
Welfare gain	CHF/d	0.02	0.06	0.03	0.12	0.00	0.20
		[0.02, 0.03]	[0.05, 0.06]	[0.03, 0.03]	[0.12, 0.13]	[-0.00, 0.00]	[0.19, 0.21]
Welfare gain per year	CHF/y	8.6	20.8	10.3	45.1	0.0	72.2
		[8.04, 9.15]	[19.62, 22.07]	[9.72, 10.96]	[42.64, 47.75]	[-0.26, 0.33]	[68.46, 76.29]
Share of baseline welfare gain	%	35	86	42	186	0	298

Notes: The first three columns impose different pricing but the same valuation of external costs as in our base setting. Specifically, CO_2 tax is equal to the Pigovian climate externality, Health tax is the Pigovian health externality applied to car and public transport, and the perimeter tax is the mean congestion externality per kilometer applied to car and public transport distances within the perimeter, respectively. In contrast, the last three columns use the same pricing but different external cost parameters. The new SCC is CHF 430/tCO_2. In "Ext. costs halved" and "doubled", the true external costs are assumed to be 50% and 200%, respectively, of the baseline.

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 0.38 corrections per person-day, and at least one correction on 20% of the person-days. As the mode is crucial in determining the external costs, the possibility of overwriting the detected mode for a particular stage 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 mode mis-assigned by the app. 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. The first two columns 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.[32]

Table 10: Corrections and spatial discontinuities

	Nr. of corrections	Correction (1/0)	Nr. of jumps	Jump (1/0)
Pricing	-0.044	-0.005	0.057	0.002
	(0.042)	(0.006)	(0.073)	(0.002)
Information	-0.056	-0.008	-0.077	-0.003
	(0.037)	(0.005)	(0.063)	(0.002)
N	112,741	168,359	107,992	168,359

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 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 to identify significant spatial jumps.[33] We record an average of 0.04 jumps per person-day. The last two columns 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 relative to the baseline. We can furthermore 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). Tables B.6 and B.7 show that these distances are very similar, and that the ATE on distance by model is essentially unchanged. Taken together, these robustness tests imply that our results are unlikely to be driven by strategic mode correction.

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

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

Table 11: Robustness tests

	Baseline	w/o corrections	w/o zeros	w/o weeks 5-6	w/o weeks 7-8
Pricing	-0.230**	-0.231**	-0.228**	-0.247**	-0.247**
	(0.070)	(0.077)	(0.072)	(0.080)	(0.083)
Information	-0.098	-0.093	-0.094	-0.090	-0.102
	(0.067)	(0.074)	(0.069)	(0.079)	(0.081)
heat	-0.498**	-0.528**	-0.559**	-0.483**	-0.487**
	(0.073)	(0.082)	(0.073)	(0.075)	(0.074)
cold	0.188**	0.182**	0.203**	0.181**	0.197**
	(0.018)	(0.020)	(0.019)	(0.021)	(0.021)
rain	0.001	0.000	0.004	0.002	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)
Adj. R^2	0.234	0.238	0.239	0.237	0.233
Clusters	3,616	3,616	3,616	3,616	3,616
N	168,359	137,896	160,947	126,058	128,156

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 overall very high in our sample, many participants did not record positive distances on all days. To differentiate between true zeros (i.e., participants not traveling) and missing data (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”) lasting 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. We treat such person-days as missings rather than zeros.

However, the imputation of activities and locations is not always correct. To assess 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 and is very similar to the baseline.

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 to nonrandom attrition should increase over the course of the treatment phase. We therefore re-estimate our base model using only the first two weeks or the last two weeks of the treatment period (columns 3-4 of Table 11). 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).[34] The results show that observability correlates positively with pre-treatment external costs (column 1 in Table B.13). However, when conditioning on pre-treatment observability (as we do implicitly in our analysis 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 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

[34] 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.

conditions, namely selection, attrition, naturalness and scaling. 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 "tracking study" 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.

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) and longitudinal scaling. 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. 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 a mode shift away from driving. This requires that public transport be available, and 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 the 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, Thaler and Johnson (1990) show that individuals tend to combine prior gains with subsequent losses, which facilitates risk-seeking behavior until the prior gain is completely depleted. Since our participants start with a gain in the form of a 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 to January in Switzerland. Although this includes a number of weeks with relatively mild climate, the colder part of the year slightly 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.

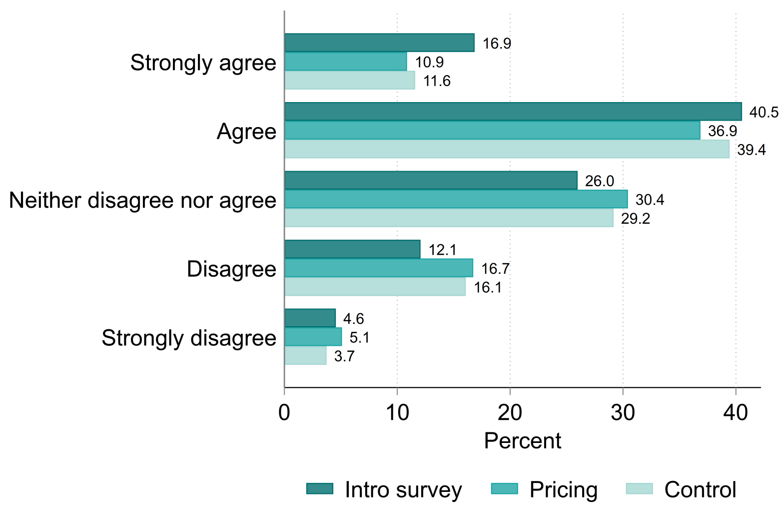
Most importantly, the treatment period in our experiment lasted only one month, such that we can only measure short-term responses. With a permanent introduction of transport pricing, additional margins of response will become available such as the choice of work and home locations, changes in activity 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). Future studies are needed to better understand the implications of externalities-based transport pricing in the long run, and in different settings.

6.5 Social acceptance

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.

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 introduction and final surveys about the extent to which they agreed with the following statement: "The price for mobility should reflect the social cost (e.g., health, environment, congestion)." Figure 8 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.[35] We do not find a significant difference across the treatment groups.

Figure 8: 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),1,070 (control); 1,058 (pricing).

[35] 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).

7 Conclusion

The MOBIS experiment implemented Pigou-inspired transport pricing based on sample tracked by GPS. The short-term treatment effect consists in a reduction of transport-related external costs by 4.6%. This 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. 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), and our results are consistent with a subsequent experiment in Oslo that uses a similar approach to ours (Ciccone et al., 2025). Whereas the information-only treatment had a statistically significant effect for a subsample that we identify as above-average "altruistic", there is no clear effect for the sample overall. Our results therefore imply that while the information content of transport prices may play a role the overall response, information campaigns will not be sufficient to induce significant behavioral change.[36] Information plays the smallest role in terms of reducing congestion externalities (the dimension for which the "pricing only" treatment is statistically significant), possibly because the high internal cost makes the congestion externality more salient than other external costs.

Our experiment shows that multi-modal transport pricing works in practice. The required technology is available, and could be implemented in principle. Pigovian transport pricing is an alternative funding mechanism that can also be implemented in the presence of a sizeable electric vehicle fleet. However, a Pigovian pricing scheme thar relies on continuous tracking 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. However, even a simplified pricing scheme should be guided by the marginal external costs of transport to increase the efficiency of the transport system. We find that a fuel and fare tax calibrated to capture the climate and health-related external costs of driving and public transport would achieve 86% of the welfare gain of a pricing that varies by time and space. Considering the implementation challenges and privacy considerations, increasing the fuel tax to incorporate (most of) the external costs of transport may be preferable. To incorporate the expansion of the electric fleet, fuel taxes would have to be accompanied by a volumetric tax on EVs.

However, the degree to which fuel taxes can capture the welfare benefits from a full set of

[36]Our point estimate of -0.1 CHF/day associated with the information-only treatment has a p-value of 0.145. Using -0.1 as the minimally detectable effect, power calculations in the spirit of Burlig et al. (2020) reveal a power of only 18% to detect an effect at $p<0.05$. It is therefore not clear if no information-only effect exists, or if our experiment was simply under-powered to detect it.

transport prices will depend on the setting. In Switzerland, congestion costs contribute only around 20% of the transport externalities. In places where congestion accounts for a larger share of the unpriced external costs (e.g., London or New York City), the benefits of pricing that (also) targets congestion will arguably be greater. Exclusively focusing on congestion is likely inefficient too due to the external costs that are unrelated to capacity constraints.

Independent of the exact shape of the pricing, 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). Using a price that is very far from the Pigovian rate may negate any welfare benefits from such intervention.

The transport pricing implemented in our experiment is regressive, a feature that it shares with fuel taxes (West and Williams, 2004; Bento et al., 2009). In order to protect poorer households and to win political acceptance, any efforts to advance transport pricing will need to be complemented with re-distributive measures to counteract adverse distributional implications. However, if implemented in an equitable way, transport pricing could become a key pillar of sustainable transport policy.

Data availability statement

The replication files are available at Zenodo `https://doi.org/10.5281/zenodo.17965910`.

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.

Anderson, Michael L and Maximilian Auffhammer (2014). "Pounds that kill: The external costs of vehicle weight." *Review of Economic Studies* 81(2): 535–571.

Andor, Mark A, Andreas Gerster, Kenneth T Gillingham and Marco Horvath (2020). "Running a car costs much more than people think—stalling the uptake of green travel." *Nature* 580(7804): 453–455.

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).

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

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.

Ciccone, Alice, Cloé Garnache and Gøril L Andreassen (2025). “Road Pricing: Travel Behavior and Public Support.” *CESifo Working Paper No. 11867* .

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 Chaisemartin, Clément and Xavier d’Haultfoeuille (2020). “Two-way fixed effects estimators with heterogeneous treatment effects.” *American Economic Review* 110(9): 2964–2996.

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.

Ecoplan / Infras (2024). “Externe Effekte des Verkehrs 2021. Umwelt-, Unfall-und Gesundheitseffekte des Strassen-, Schienen-, Luft-und Schiffsverkehrs.” technical report, Bern. `https://www.are.admin.ch/dam/are/de/dokumente/verkehr/dokumente/bericht/GKV21_SB_final_inkl_%C3%9Cbersetzung_KF.pdf.download.pdf/GKV21_SB_final_inkl_%C3%9Cbersetzung_KF.pdf`.

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.

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.

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 (2024). "Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications." *Econometrica* 92(4): 1233–1268.

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.

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.

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

Swiss Federal Office of Spatial Development (2024). “Analysis of the Stated Preference Survey 2021 on Mode, Route and Departure Time Choices.”Technical report, Swiss Federal Office of Spatial Development (ARE), Berne.

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.