THE IMPACT OF SOCIAL EXTERNALITY INFORMATION ON FOSTERING SUSTAINABLE TRAVEL MODE CHOICE: A BEHAVIORAL EXPERIMENT IN ZHENGZHOU, CHINA

Rachel Luo[1]
rlluo@mit.edu
Master's Candidate in City Planning and Transportation

Yichun Fan[1]
ycfan@mit.edu
Ph.D Candidate in City Planning

Xin Yang[2]
xinyang@smart.mit.edu
Postdoctoral Associate, Singapore-MIT Alliance for Research and Technology (SMART)

Jinhua Zhao[3]
jinhua@mit.edu *corresponding author
Edward H. and Joyce Linde Associate Professor of Transportation and City Planning
ORCiD: 0000-0002-1929-7583

Siqi Zheng[1]
sqzheng@mit.edu
Samuel Tak Lee Professor
Faculty Director, Sustainable Urbanization Lab
Faculty Director, MIT Center for Real Estate
ORCiD: 0000-0002-4467-8505

[1] Department of Urban Studies and Planning, Massachusetts Institute of Technology
77 Massachusetts Ave., Cambridge, MA 02139, United States
[2] SMART, 1 CREATE Way, #09-01/02 CREATE Tower, #01-13 Enterprise Wing, Singapore 138602
[3] Corresponding author. Department of Urban Studies and Planning, Massachusetts Institute of Technology. 77 Massachusetts Ave., Cambridge, MA 02139, United States

AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows—study conception and design: S.Z., J.Z., Y.F., R.L.; data collection: Y.F., S.Z.; analysis and interpretation of results: R.L., X.Y., S.Z.; draft manuscript preparation: R.L., X.Y.

ABSTRACT

Urban leaders in areas with high air pollution often face the dual task of reducing pollution levels while educating the public about the health impacts of pollution and preventive measures. Transportation policies to cut motorized personal vehicle use are often a key part of pollution reduction efforts. One type of these policies is information interventions educating commuters of the higher emissions impact of cars. This study evaluates the impact of such an information intervention on car commuters' intention to switch from car use to transit, biking, or walking in Zhengzhou, China. Further, it tracks how this greening impact evolves as drivers are given additional public health information regarding Zhengzhou's severe air pollution level, its health effects, and the reduced pollution exposure when driving compared with transiting, biking, or walking outside in polluted air.

Using a randomized controlled survey experiment with a treatment group of 281 participants and a control group of 280 participants drawn from Zhengzhou workers who typically commute by car, this study applies multinomial logit models and difference-in-difference estimation to estimate the information intervention's impact. We find that while emissions information interventions are initially highly effective in reducing intended car usage, the effect is progressively dampened as respondents learn of the level of local air pollution and the risks of outdoor pollution exposure. Some greening effects do, however, remain even at the end of the experiment. The results suggest that governments in polluted areas must be aware of how policies in different segments of their portfolio interact and support broader initiatives like mixed-use planning and transit extension, simultaneously cutting transportation emissions and creating shorter commute durations with lower outdoor pollution exposure.

Keywords: Air pollution, behavioral intention to change, information intervention, transportation mode choice, green mobility

1. Introduction

Well-documented links between air pollution and public health problems have prompted governments around the globe to take actions to reduce air pollution, and raise public awareness of ways to limit personal pollution exposure. Epidemiological studies tie air pollution components like particulate matter ($PM_{2.5}$, PM_{10}) to higher incidences of lung disease (Raaschou-Nielsen, et al. 2013, Wang, et al. 2017), cardiovascular disease (Brook, et al. 2010, Chin 2015), and lower worker productivity (Zivin and Neidell 2018). Overall, exposure to ambient air pollution is estimated to cause up to 3.3 million premature deaths per year globally, making it the largest environmental health risk (Hamanaka and Mutlu 2018, Lelieveld, et al. 2015). In China, air pollution has been estimated to lower life expectancy by three years (Ebenstein, Fan, et al. 2017), and $PM_{2.5}$ exposure is estimated to be responsible for 19-40% of four of China's five top causes of death (Song, et al. 2017).

Vehicle traffic is a leading cause of air pollution, especially particulate matter, in many regions including China. Transportation reform is thus a key lever for governments to curtail air pollution and its associated health consequences. In our study area, the mid-size city of Zhengzhou in central China, transportation contributes an estimated 32% of $PM_{2.5}$ (Gong, et al. 2017). Research has also established direct links from traffic exposure to lung cancer (Raaschou-Nielsen, et al. 2013) and heart disease (Nawrot, et al. 2011), exposing the importance of measures to reduce motor vehicle traffic. Further, another vehicular pollutant, greenhouse gases, have long been a target of government regulation due to its effect on climate change (Hensher 2008, Woodcock, et al. 2009).

Policymakers around the world have therefore tried to incentivize a public shift away from private car usage to less polluting modes like public transit, and "active" transportation modes such as biking and walking. Such policies include traditional methods like infrastructure expansion, congestion charges, license plate restrictions, and transit pass subsidies, but also "information interventions" targeting behavioral change (Geng, Long and Chen 2016, Kishimoto, et al. 2017, Waygood and Avineri 2018).

"Information interventions" here indicate education and awareness efforts targeting individual behavior change by giving relevant knowledge regarding the costs and benefits of the alternatives under consideration. For example, past research has examined the effect of information regarding modal carbon emissions (Waygood and Avineri 2016), air pollution (Mir, et al. 2016, Saberian, et al. 2017), public health impacts (Mir, et al. 2016), or a combination of these (Geng, Long and Chen 2016). Others have given accessibility information (Guo and Peeta 2019) or suggested providing travelers with travel time and cost information for cleaner alternative modes (Löfgren and Nordblom 2006). This study focuses on carbon emissions information as a way to induce individuals to switch away from commuting by car, through showing them the social externalities of their car use.

Concurrent with transportation management, urban governments are building public awareness of air pollution, its consequences, and preventive measures to reduce personal pollution exposure. This is another type of information intervention, directly targeting public health. In China, household awareness of local pollution levels rose in the past decade with the rollout of air quality sensors and mobile app access to real-time pollution data, leading residents to adopt preventive measures including higher willingness-to-pay for residences in less polluted areas, limiting time outdoors, using air purifiers, and wearing face masks (Barwick, et al. 2019, Chen, Oliva and Zhang 2017, Sun, Kahn and Zheng 2017). Barwick et al. (2019) estimated that the health benefits of releasing this pollution information is an order of magnitude larger than the cost of program deployment plus the cost of taking the preventive actions.

Yet for urban mobility, preventive actions to reduce air pollution exposure, such as avoiding outdoors time and looking for air-purified areas, can perversely increase the usage of polluting private vehicles rather than cleaner active or transit modes. Pollution exposure varies depending on mode and an area's ambient pollution level (Cepeda, et al. 2017, Tainio, et al. 2016). In highly polluted locations such as Zhengzhou, however, air pollution exposure from the increased time spent outdoors while walking, biking, or accessing transit is a more visible health concern for residents. Past studies suggest that Chinese commuters find air pollution a barrier towards travel modes that require more outdoor time, such as biking, walking, and transit (Jiang, et al. 2017, Li and Kamargianni 2017, Yang and Zacharias 2016).

Hence, beneficial public health information regarding pollution levels may counteract transportation-sector information interventions promoting active modes and transit. This paper brings together these two key policy branches to assess 1) how information interventions conveying the differential emissions impact of travel modes affect drivers' intention to switch to non-car modes, and 2) how this effect interacts with equally essential public health information interventions in highly polluted places like Zhengzhou, China. The remainder of this paper is organized as follows: Section 2 presents relevant literature; Section 3 explains our methodology; Section 4 details our sample characteristics; Section 5 presents key results; and Section 6 concludes the paper.

2. Literature Review

Previous studies have used stated preferences or self-reported behavior to demonstrate that information interventions can promote greener travel mode decisions. Among those on interventions involving carbon emissions information, Waygood and Avineri (2016) found that travelers, including drivers, reported higher intention to reduce car use when given information on carbon emissions from car travel. The paper explored the impact of information wording on intended car use, and confirmed the treatment's effect was greater when the information was framed in a relatable way, in terms of tree-equivalents or a typical person's carbon budget. Responses were stronger for participants in middle-income countries like China. The results were based on logit models and data from surveys distributed at work or in focus group in five countries.

Building on that line of work, Waygood and Avineri (2018) informed individuals of their neighborhood's carbon emissions from travel and compared their performance against an achievable lower emissions rate. This comparative framing leverages the power of social norms and was associated with greater respondent motivation to reduce travel emissions through switching to green modes and reducing the amount of travel. Their results were based on logit models applied to data from an online survey that randomized respondents into experimental groups with different emissions framings. Moreover, for car purchases, Daziano et al. (2017) presented respondents with carbon emissions intensity information in a discrete choice experiment and found, using multinomial logit models, that the information boosted respondents' willingness to pay for lower-emission vehicles.

As the Methods section below will describe, this paper leverages some of the methods used in the above literature and shares their focus on carbon emissions information. However, while the literature focuses on improving the psychological framing of information, our work explores how the impact of emissions information on intended car usage changes when individuals encounter equally important public health information that recommends a countervailing shift in mode choice for self-protection against pollution exposure.

Other information intervention studies do, like us, include a broader array of information beyond carbon emissions. Mir et al. (2016) found that general information on the negative health and environmental effects of air pollution and carbon dioxide, especially when framed in terms of the gains achievable through pollution mitigation, shifted individuals' intended travel mode toward more sustainable options in Tehran, Iran. The study applied ordinal regressions to survey data from 220 students in Tehran, who were randomized into four groups with different pollution impact and location framings. Geng, Long, and Chen (2016) examined the cumulative impact of a multi-faceted information intervention on car use time, focusing on how impact varied between people with different motivations, or "goal frames." The information intervention included the health and convenience benefits of using active modes or transit; the personal financial cost of car ownership; and the health and emissions costs of driving. Delivered through a randomized controlled experimental design, the information consistently increased self-reported travel time spent on green modes compared to the baseline. Yet, difference-in-difference analysis (DID) showed the decline in car use time was insignificant, and effect sizes varied by goal frame.

While these previous studies estimated the aggregate effect of a bundle of related information, we provide a more granular look at the effect of modal carbon emission information on intended car usage and how that evolves with each additional type of public health information given. Morever, our information interventions are different in that they are personalized, with carbon emissions and pollution exposure provided for every modal alternative tailored to the respondent's commute. These design differences arise because our goal is to help policymakers understand interactions among elements of a policy portfolio, while existing literature on information interventions focus solely on respondent psychology and information wording for a particular dimension of information. In this, we follow the lead of some research outside of the information intervention domain, which has attempted to assess interactions between different green mobility policies (Kishimoto, et al. 2017).

Considering the interaction of public health information with emissions information is of real-world value since the results from Mir et al (2016) and Geng, Long, and Chen (2016) suggest pollution level and health data can trigger change in intended mode use. In addition, many public health studies have suggested that personal pollution exposure information affects traveler mode choice among other preventive actions (Jiang, et al. 2017, Li and Kamargianni 2017, Yang and Zacharias 2016), which lend some support to our study's general premise.

In summary, drawing from the literature, this paper integrates randomized controlled trials in survey designs to evaluate the effectiveness of carbon emissions information interventions in fostering commuters' intention to reduce car use. Based on existing work, our emissions information is worded to help respondents contextualize the size of their emissions and compare alternative modes. Further, this study extends the literature by examining the complex reality of interacting interventions within a government's policy portfolio, an area noted as a policy analysis gap in China (Jiang, et al. 2017). It does so with multinomial logit models in the vein of preceding papers, plus DID analysis enabled by our randomized controlled experimental design. Multinomial logit models provide insights into the net outcome after each round of intervention, while DID allows us to examine divergence in the rate of change in car use intention between the control and treatment groups as they receive new information throughout the experiment.

3. Methods

3.1 Experiment design

This paper is based on a stated preference survey executed in Zhengzhou, China in July 2019. The survey was built in Qualtrics, a professional survey tool, and distributed to company employees working in the Zhengzhou Central Business District (CBD). Recruiting respondents from workplaces has precedent in other information intervention studies (Waygood and Avineri 2016). Since this experiment focused on shifting car commuters toward public transit or active travel modes, it was administered to 607 survey respondents who, in a pre-screening, declared cars as their primary commuting mode (18% of pre-screening responses). From these, we received 561 valid, complete responses. More details about our procedure are available in Appendix A.1.

Respondents were randomized into the treatment or control track. Using a control group enables us to adjust our analysis for systematic influences biasing the overall behavior of survey participants, to the extent that these factors affect the control and treatment group in similar ways. This allows us, for example, to mitigate the experimenter demand effect where respondents try to appear socially responsible when they perceive that researchers deem this desirable behavior (Grimm 2010). Further, it mitigates the common survey design concern that respondents' answers to later survey questions may be affected by their responses to earlier rounds of questions. Individual travel behavior is also derived from a complex array of influences, so comparing the treatment group against a statistically indistinguishable control group allows us to make a cleaner identification of the effect of the carbon emissions information intervention (i.e., the treatment).

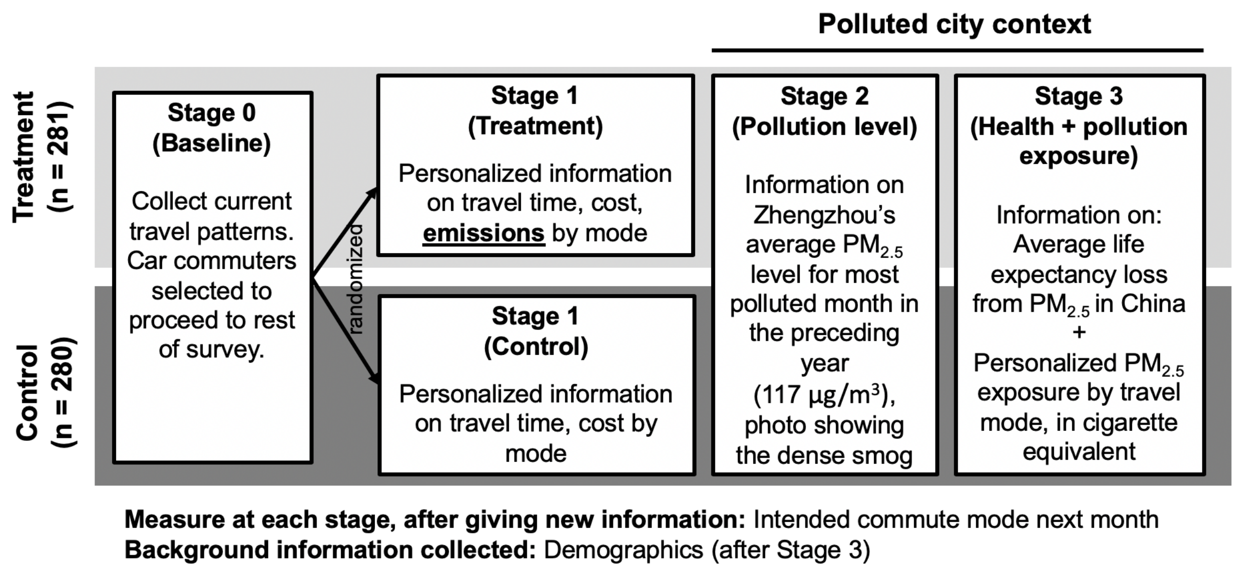


Figure 1. Stated preference randomized controlled experiment design

The survey applied a series of information interventions to both the control (n = 280) and treatment (n = 281) groups, but only treatment group participants received an information table listing how carbon emissions differed by mode for their personal commutes (Figure 1). This carbon emissions information is called the "social externality information intervention." This study evaluates the impact of the social externality information intervention on intended car use, and how the impact evolves as respondents received additional pollution and health information. Background information on respondent demographics were collected from both groups as control variables for analysis. The questionnaire for our relevant respondents can be found on Github: https://github.com/rlluo1/Zhengzhou_2019_survey[1]

Both groups began in baseline *Stage 0*, when respondents inputted their existing commute patterns including the travel mode. This question directly asked respondents for their typical daily commuting mode (Appendix A.2) and was used to pre-screen for car commuters, our survey's target demographic. Car commuters were then randomized into the control or treatment group.

In *Stage 1*, both the control and treatment group respondents inputted their home and work addresses into a customized website that used the API of Amap, a major Chinese map service provider, to inform respondents of their personalized travel time by car, public transit, walking, or biking, as well as their personalized travel cost by public transit or taxi. This provided all participants with information on their travel alternatives, giving them a factual basis for modal comparison (Gaker and Walker 2013, Waygood and Avineri 2018). The treatment group members received additional rows in the table showing personalized greenhouse gas emissions by mode. To help respondents contextualize the emissions information, the survey presented this data in terms of the number of trees needed to absorb these annual emissions, based on one of the best-performing formats from Waygood and Avineri (2016)[2]. Neither Stage 0 nor Stage 1 included any information on Zhengzhou's pollution level or its health impacts; further, the survey was given in summer, when the pollution level was much lower ($PM_{2.5}$ under 40 μg/m^3).

In *Stage 2*, respondents were informed that Zhengzhou's local $PM_{2.5}$ level averaged 117 μg/m^3 during the most polluted month in the preceding year, accompanied by a photograph capturing the dense smog in the city (Appendix A.4). In *Stage 3*, respondents were informed of University of Chicago research that found air pollution decreased average life expectancy by three years in northern China (Ebenstein, Fan, et al. 2017, Ebenstein, Fan, et al. 2015). An accompanying table showed respondents their personalized commute-time $PM_{2.5}$ pollution exposure estimated by mode. This table illustrated that driving in properly air-filtered cars protects commuters from pollution exposure, because the traveler can avoid walking outside to bike, walk, or access transit.

Because air pollution exposures vary widely, our research team built this modal pollution exposure table by combining existing literature with on-site data we collected over two weeks using four professional air pollution monitoring devices from Fairsense.[3] Twice per day for two weeks, team members commuted during peak hours by bus, subway, car, bicycle, and walking along three major commuting routes around the respondents' workplaces in Zhengzhou CBD, recording pollution measurements along the routes. Combining this local pollution severity data with modal differences in inhalation rate (i.e., how much polluted air commuters inhale when using each mode) from literature (Cepeda, et al. 2017) allowed us to provide respondents with information more representative of Zhengzhou local conditions (Appendix A.3). The exposure intensity data were then multiplied by each respondent's real commute time by mode to obtain a personal modal pollution exposure value. Further, to help respondents contextualize their pollution exposure, we presented the monthly exposure

[1] The questionnaire is posted in English to match the language of this manuscript. However, the original survey was conducted in Mandarin Chinese, as shown by the example in Appendix A.4

[2] The key information interventions were presented in tables and bar charts, and surveyors were recruited and trained to present the information in a neutral way to survey takers. These efforts were made to ensure participants properly understood the presented information.

[3] Fairsense Environmental Technology is a technology company specializing in environmental monitoring and environmental research. Details can be found at http://www.fairsense.cn/

values in terms of cigarette equivalent; exposure to $PM_{2.5}$ levels of 22 μg/m^3 for one day translates approximately to smoking one cigarette (Muller and Muller 2015).

At the end of each stage, respondents were asked in neutral language for their intended commute mode next month, considering all the information given. The wording is similar to the corresponding baseline question, and provides the outcome variable used in this paper (details in Appendix A.2 and on GitHub).

All experimental stages provided information interventions of some type—on mode alternative characteristics, local pollution severity, and pollution exposure by mode plus pollution's mortality effect. However, only Stage 1 differed between the control and treatment groups. Comparing the two groups thus allows us to isolate the effect of Stage 1's social externality information intervention. Note also that the social externality information used carbon emissions, which is an often-used method in the papers from our literature review. Meanwhile, Stages 2 and 3 provided public health information on the more visible and toxic $PM_{2.5}$; this was meant to reflect Chinese national air quality reporting, shown by recent studies to have triggered preventive action (Barwick, et al. 2019, Chen, Oliva and Zhang 2017, Tu, et al. 2020). The survey thus did not target one pollutant, but engaged respondents generally about transportation-related pollutants.

The data is limited by the fact that aside from in baseline Stage 0, the response variable collected was intention to switch, not actual switching behavior. This was because participants were only recruited for one 15- to 20-minute period to take a survey near their workplace; we did not create an app through which to deliver or collect further information as Geng *et al.* (2016) did. Intention and stated preferences, though not ideal, are commonly used in this vein of research due to the difficulty of tracking individuals long-term (Gaker and Walker 2013, Heinen 2016, Mir, et al. 2016, Waygood and Avineri 2016, Waygood and Avineri 2018). Further, under theories like the Transtheoretical Model (Prochascka, Redding and Evers 2008) and the Theory of Planned Behavior (Ajzen 1991), intention is an intermediate step to actualized action. Lastly, the randomized controlled experiment design used here allows us to control for survey biases that affect both the treatment and control groups to the same degree.

3.2 Data analysis methods

Analysis was conducted with multinomial logit (MNL) models and difference-in-difference (DID). MNL models were applied to intended commute mode choice responses from Stage 1, 2, and 3, in order to understand the *level* of intended car-use responses stated at each stage, and associate this with the treatment while controlling for individual demographics and alternative-specific variables (Hensher, Rose, and Greene 2015). Then, an individual-level DID regression examined how the *change* from stage to stage in intended car use varied between the treatment and control groups.

Logit models have long been used for choice modeling, and have also appeared in many information intervention papers (Daziano, et al. 2017, Gaker and Walker 2013, Waygood and Avineri 2016, Waygood and Avineri, 2018). Because the walking share is small in each stage, we condensed the response variable to three categories: car, transit, and active travel mode (i.e. biking and walking). The probability that a respondent *i* chooses mode *j* at any one stage is predicted by:

$$P_i(n = j) = \frac{e^{V_{i,j}}}{e^{V_{i,car}} + e^{V_{i,transit}} + e^{V_{i,active}}} \quad \text{Eq. 1}$$

In the above expression, the three deterministic utilities $V_{i,car}$, $V_{i,transit}$, and $V_{i,active}$ are specified as:

$$V_{i,car} = \beta_{time,car} \ln(time_{i,car}) + \sum_{q=1}^{s} \beta_q z_{iq} \quad \text{Eq. 2}$$

$$V_{i,transit} = \beta_{transit} + \beta_{treated} treated_i + \beta_{time,transit} \ln(time_{i,transit}) + \beta_{cost} cost_{i,transit} + \sum_{m=1}^{k} \beta_m x_{im} + \sum_{q=1}^{s} \beta_q z_{iq} \quad \text{Eq. 3}$$

$$V_{i,active} = \beta_{active} + \beta_{treated} treated_i + \beta_{time,active} \ln(time_{i,active}) + \sum_{m=1}^{k} \beta_m x_{im} + \sum_{q=1}^{s} \beta_q z_{iq} \quad \text{Eq. 4}$$

where,

$\sum_{m=1}^{k} \beta_m x_{im}$ represents the *k* = *9* sociodemographic variables described below,

$$\sum_{m=1}^{k} \beta_m x_{im} = \beta_{ownscar} ownscar_i + \beta_{errands} errands_i + \beta_{male} male_i + \beta_{hhsize} hhsize_i + \beta_{age} \ln(age_i) + \beta_{local} local_i + \beta_{ownshome} ownshome_i + \beta_{income} income_i + \beta_{college} college_i \quad \text{Eq. 5}$$

and $\sum_{q=1}^{s} \beta_q z_{iq}$ represents *s* interaction terms between the alternative-specific travel time/cost variables with sociodemographics.

The above equations are estimated separately for each experimental stage. $\beta_{transit}$ and β_{active} are alternative-specific constants (ASCs); *treated* is an indicator for whether individual *i* is in the treatment group; *time* is travel time for each mode; and *cost* is modal trip cost. Travel times and age were logged to correct for right skew. Car is the reference category, so its specification did not include an alternative-specific constant or socio-demographics. Positive coefficients in the transit and active mode equations should thus be interpreted as indicating greater likelihood that the traveler will choose that mode instead of car.

The treatment effect is the key variable of interest here. Other variables were included as controls because they are commonly shown in the literature to explain mode choice, and because they are meaningful in the context of our analysis. They belong to three categories:

- ***Socio-demographics:*** This includes the number of cars owned by the respondent's household, whether the person is a Zhengzhou local, and the person's gender, household size, age, homeownership, income group, and college degree. We also include whether the respondent regularly runs errands during commutes. Household size, errands, and age are commonly included in mode choice modelling. Car ownership is included as it enables commuting by car, in China and globally (He and Thogersen 2017). Income, education, and homeownership capture relationships between car use and wealth or status. Driving and car ownership are status symbols in China, suggesting the wealthy may be less likely to give up commuting by car due to social signaling (Zhao and Zhao 2020, He and Thogersen 2017). Gender is included because studies suggest women, more than men, perceive pollution as unsustainable (Waygood and Avineri 2018). Lastly, though Mir et al (2016) suggests psychological distance is not a significant factor in this type of study for Tehran, we include an indicator of whether the respondent is a Zhengzhou native to test if in China, locals are more reactive to information about their home environment.

- ***Alternative-specific characteristics:*** Variables capturing travel time and cost by mode were presented to respondents in Stage 1 and are included in the model, as they are typical in travel mode choice modelling (Ben-Akiva and Lerman 1985). However, our customized website did not present the marginal cost of commuting by driving a personal car (e.g. fuel cost), only taxi/ridehail costs which were less relevant to our respondents, most of whom drove personal cars.[4] Therefore, one weakness of our data is that it did not enable us to include personal car commuting costs in the relevant utility function. Additionally, the website did not explicitly indicate the availability of transit or active alternatives for each respondent. We assumed active modes were available to respondents (walking or biking; bikeshare is widespread in Zhengzhou[5]). The best possible car, transit, and active mode options determined by Amap were all offered at each stage to every respondent. Therefore, we did not limit the availability of any of the modes in the choice set. However, the Amap API showed realistic travel times for each mode during the respondent's typical commute time for their typical origin-destination pair; this meant displaying substantially longer travel times for transit where it is less directly accessible. Actual transit costs were also included, though there is limited variability in this figure across Zhengzhou due to the largely flat fare system. So, respondents were able to make their mode choice decision with a realistic sense of the time and cost commitment for each alternative.
- ***Interaction terms between travel time/cost and sociodemographics:*** These capture possible systematic heterogeneity in respondent preferences for travel time and cost. Subsets of interaction terms are selected for inclusion in the model from the broader pool of all possible interactions, when doing so improves the model fit.

Next, individual-level DID regression is performed to analyze how the treatment influenced changes in car use intention from stage to stage. Performing DID as linear regression[6] allowed us to separate the effects common to both the control and treatment groups across the experimental stages ("time effects"), from the effects specific to the social externality information treatment at each stage ("DID effects"). DID regression also allows for the inclusion of time-invariant control variables (Dimick and Ryan 2014). The equation below specifies our regression, where the outcome variable *Y* is a binary value indicating whether an individual *i* intends to commute by car (Y = 1) or not (Y = 0) during a particular experimental stage *t*:

$$\begin{aligned} Y_{it} = {} & \beta_0 + \beta_1\ treated_i + \beta_2\ Stage_t + \beta_3(treated_i * Stage_t) + \\ & \beta_4 \ln\left(car\_time_{\ i}\right) + \beta_5 \ln(transit\_time_i) + \beta_6 \ln\,(active\_time_i) + \ \beta_7\ (transit\_cost_i) + \\ & \textstyle\sum_{m=1}^{k} \beta_m x_{im} + \tau_t + \varepsilon_{it} \end{aligned} \qquad \text{Eq. 6}$$

The regression above is conducted three times, once for each experimental stage (that is, t = 1,2,3). The first assesses the DID effect between treatment and control groups from baseline Stage 0 to Stage 1, the second from Stage 1 to Stage 2, and the third from Stage 2 to Stage 3. The DID effect of each stage is captured by β_3 in our regressions.

---

[4] For details, see footnote 8 in Section 5: Results.

[5] Bike accessibility is high in Zhengzhou. The local government sponsored free docked public bikes for residents and there are many private dockless bikesharing programs providing cheap biking services.

[6] Though the outcome is binary, we use linear regression with robust standard errors instead of logit or probit. Linear regression is limited here because it does not bound the response variable to between zero and one. However, because difference-in-difference is an interaction effect, it cannot be directly translated into non-linear forms such as probit and logit; doing so also fails to meet DID's common trends assumption. Methods for addressing this face difficulties of interpretation and estimation, so we use linear regression instead (Ai and Norton 2003, Karaca-Mandic, Norton and Dowd 2011, Lechner 2010, Puhani 2012).

The *treated*$_i$ variable represents an individual *i*'s membership in the treatment group and captures any treatment effect that stays constant across experimental stages. The *Stage*$_t$ variable in each equation captures the "time effect" common to both the control and treatment groups as they pass through experimental stage *t*; these include information received by both groups such as the modal alternative characteristics given in Stage 1, the local pollution level given in Stage 2, and the pollution exposure by mode in Stage 3).

The interaction of each time effect with *treated*$_i$ represents the DID effects, whose coefficients β_3 are the key point of interest for this portion of the analysis. The coefficients for these DID interaction terms capture differences in the rate at which treatment and control groups changed their intention to commute by car as they progressed through the experimental stages. Finally, as in the MNL, $\sum_{m=1}^{k}\beta_m x_{im}$ represents the $k = 9$ sociodemographic variables described in Eq. 5; we also again include the alternative-specific variables of travel time and (for transit) travel cost.[7]

4. Sample Characteristics

Randomization between the treatment and control groups appear successful. T-tests, proportion tests and chi-squared tests did not detect significant differences between the control and treatment groups at the α=0.05 significance level across relevant sociodemographic and alternative-specific variables (Table 2). This provides some assurance with regards to our ability to validly compare the treatment against the control group to measure the treatment effect.

Table 2: Sample characteristics of control and treatment groups

Variable	Treatment (n = 281)	Control (n = 280)	P-value
Age (mean)	32.3	32.1	0.67
Household car ownership (mean)	1.59	1.54	0.45
College education or above	65%	66%	0.81
Errands during commute	18%	18%	0.99
Household size (mean)	3.8	3.7	0.24
Owns a home	69%	63%	0.20
Income group (annual)			0.78
Under RMB 50,000	8.5%	8.2%	
RMB 50,000 – 150,000	42%	45%	
RMB 150,000 – 300,000	34%	34%	
RMB 300,000+	15%	13%	
Zhengzhou local	45%	40%	0.28
Male	43%	43%	0.99
Car commute time (mean)	27 minutes	27 minutes	0.77
Transit commute time (mean)	62 minutes	61 minutes	0.45
Active commute time (mean)[xx]	52 minutes	49 minutes	0.44
Transit cost (mean)	RMB 2.20	RMB 2.30	0.47

Note: [xx] The active mode category includes both walking and biking, since walking was not often chosen. Therefore, the active commute time is that of biking for those who selected the bike mode, plus walking for those who selected the walk mode. For those who chose neither, the default was the biking time as that was the more common category. Though participants were allowed to switch between biking and walking between experimental stages, this rarely occurred. We thus only report Stage 1 active commute time here as it is representative of all stages.

[7] Additionally, the two error terms at the end of Eq. 6 capture time-specific errors and errors that occur at the individual-stage level.

The mean transit time for our sample of car commuters is twice the mean car commute time, suggesting self-selection among the car commuters—that is, the average car commuter in Zhengzhou faces relatively inconvenient transit alternatives. Assessing the distributions of the other control variables confirms that the sample is drawn from a working age population, limiting this study's applicability to students or the elderly. Perhaps because the respondents are generally younger, only a small portion report regularly running errands during their commutes (18% in both the control and treatment groups), simplifying the complexity of the commuting decisions that they face.

5. Results

5.1 Overview

This section presents results from the descriptive analysis, MNL models, and individual-level DID regression, in order. Figure 2 illustrates how intention to commute by car, public transit, or active modes evolved over the course of the experiment for the treatment versus the control groups. It suggests that at the aggregate level for all three experimental stages, the treatment group consistently reported lower intention to continue commuting by car and higher intention to use transit or active modes. However, this performance gap narrowed over time and both groups experienced a rebound in intended car use at Stage 3, when respondents were informed of the mortality effects of pollution exposure, along with the generally lower pollution exposure of commuting by car. This suggests the pollution exposure and health information reduced the effectiveness of the social externality intervention.

This descriptive result is supported by MNL outputs, which assess the *level* of intended car, transit, or active mode use at each experimental stage by using input variables to predict individual stated mode choice. The models indicate a statistically significant treatment effect after controlling for alternative-specific and sociodemographic variables, suggesting the treatment boosted the level of intention to use transit or active modes at the expense of cars for all experimental stages. Again, however, we see this treatment effect narrowing over time as pollution exposure information was introduced.

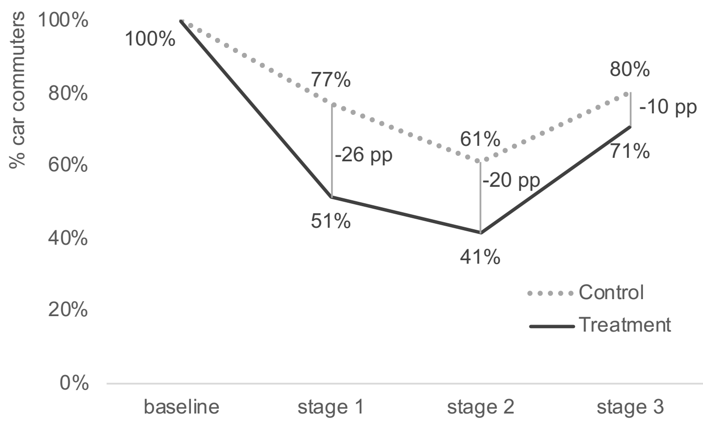
The complementary DID approach does not compare response levels directly between the control and treatment groups. Instead, it compares *changes* in the responses observed for each group from stage to stage (i.e., in Figure 2, DID tests for significant differences in the slopes of the treatment versus control group lines). By doing so, it isolates the additional change experienced by the treatment group beyond that experienced by the control group. This allows for analysis of the treatment effect when the two groups enter an experimental stage at different levels of intended car usage—as our treatment and control groups do starting at Stage 1, due to the impact of preceding information interventions. Operating on changes also avoids biases stemming from factors affecting both groups the same way over time, though it does not address biases that affect the two groups differently.

Graphical and DID results agree on some unexpected findings. While the social externality information immediately produced a steeper drop in intention to commute by car, it did not speed up the treatment group's switch away from cars once respondents were given new public health information on Zhengzhou's hazardous pollution severity in Stage 2. In Stage 3, the treatment group even rebounded toward car use more steeply than the control group did, once respondents were told that commuting by car may lower their exposure to dangerous air pollution.

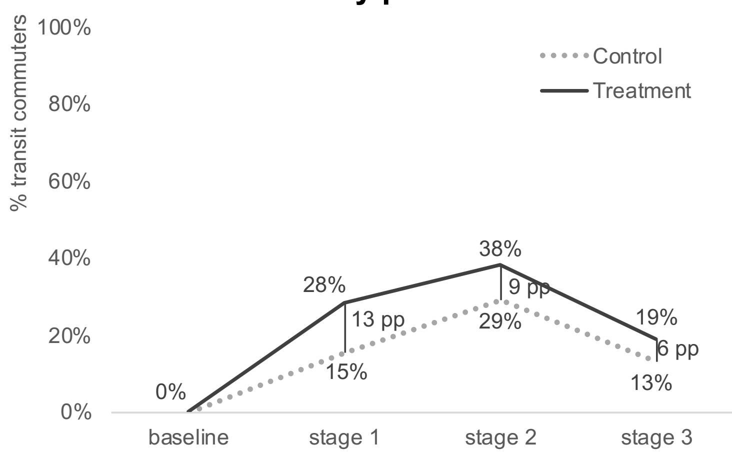
5.2 Descriptive results

Figure 2 Share of respondents intending to commute by each mode, control v. treatment group

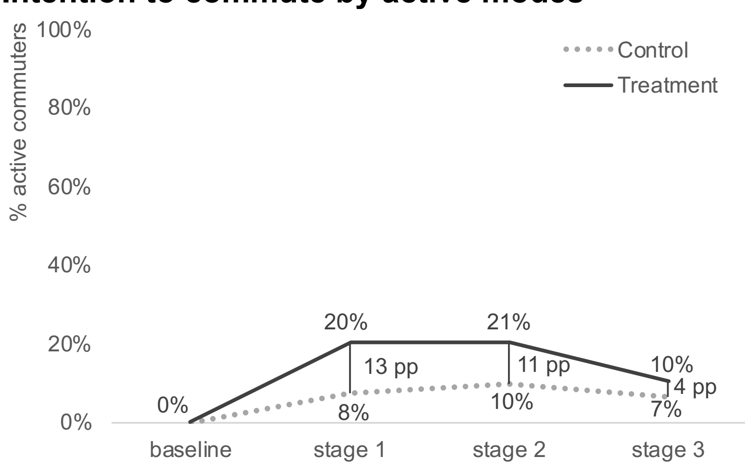
Intention to commute by car



Intention to commute by public transit



Intention to commute by active modes



The top graph in Figure 2 illustrates how intention to commute by car evolved over the course of the experiment for the control versus the treatment group. The two plots below detail how this translated to intended use of transit and active modes. Since the sample is of car commuters, 100% of respondents in both the control and the treatment group reported traveling by car in the baseline. The vertical bars labeling the difference in levels between the control and treatment groups at each stage give preliminary indication of the results of the MNL model; the differences in slopes between the two lines foreshadow the DID results.

In Stage 1, the control group only received the personalized table of commute time and cost by mode, while the treatment group received this and additional information on modal carbon emissions. Subsequently, the share of control group car commuters expressing an intention to continue using cars dropped to 77.1%, down 22.9 percentage points from the baseline. In comparison, the treatment group saw a much larger 48.8 percentage-point drop, as roughly half the group intended to try non-car modes. Most of those switching from cars intended to try transit.[8]

[8] The survey question from which we drew our response variable combined driving a personal car with riding a hired car (taxi, ridehailing) into a personal motorized vehicle category. This may raise questions regarding the degree to which the intended commute mode responses here were driven by the behavior of personal car drivers versus hired car riders. Our survey included a separate question asking the number of days per week for which the respondent typically commutes by any particular mode. Responses to this second question

In Stage 2, both the control and treatment groups received the same public health information regarding the elevated pollution in their own city. This was again followed by a drop in both groups' intended car usage, which may reflect either a sincere desire to act or a desire to appear more environmentally conscious once informed of the unhealthy level of local pollution which their own driving behaviors can exacerbate. This paper's experimental setup cannot disentangle these motivations. However, assuming any bias affects the control and treatment groups to the same extent, the randomized controlled experiment does allow us to isolate the degree to which the treatment (i.e., the high pollution scenario) triggered a differential response in the treated group versus the untreated control. The share of control group drivers intending to continue commuting by car dropped to 61.1% from 77.1% in Stage 1, a decrease of 16.1 percentage points. Meanwhile, the treatment group, though still reporting lower levels of intended car use, experienced only a 10.0 percentage point drop from 51.2% to 41.3%.

In Stage 3, as expected, drivers' intention to commute by car rebounded in both groups once they saw that using personal cars can protect them from pollution exposure. Those switching back toward cars came from both the transit and active mode groups of Stage 2. The share of car commuters intending to continue commuting by car remained slightly lower for the respondents who received social externality information, versus those in the control group. However, the rebound was much steeper for the treatment than for the control group: while the share intending to commute by car jumped 19.3 percentage points from 61.1% to 80.4% in the control group, it rose by an even larger 29.5 percentage points from 41.3% to 70.8% in the treatment group.

5.3 Multinomial logit model results

Multinomial logit (MNL) models were conducted for each experimental stage to examine the treatment effect while controlling for sociodemographic and alternative-specific factors. Table 3 summarizes the model results. The table indicates that likelihood ratio tests comparing each stage's model against a base model reflecting market share (i.e., an ASC-only model) show the variables included in the model are jointly significant in explaining commute mode selection. The higher R^2 scores on the Stage 3 and Stage 1 models compared to Stage 2 suggest that Stage 2 has a relatively weaker fit. This could be because the impact of the local pollution information provided in Stage 2 is more ambiguous without a more accurate, granular understanding of respondent attitudes. Knowing the severity of particulate matter pollution in their city, respondents may feel compelled to contribute to the social good by reducing their pollution from driving, but they may also feel compelled to select a mode that appears to reduce their pollution exposure by keeping them inside an air-filtered interior environment.

---

indicated that the vast majority of respondents in both the treatment and control groups commuted by driving personal cars. In the control group, 83% commuted by driving 4 or more days a week, and 93% commuted by driving 3 or more days a week. In the treatment group, the corresponding figures were 86% and 93%. At the same time, 77% in the control group did not regularly commute by hired cars, while 95% commuted by hired cars no more than 2 days a week. The figures for the treatment group were 81% and 97% respectively.

Table 3 Multinomial logit (MNL) model results, by experiment stage

Variable	Stage 1	Stage 2	Stage 3
Social externality treatment			
Treatment group (yes/no)	1.32 (0.201)***	1.02 (0.186)***	0.59 (0.209)***
Constants			
ASC, active mode[1]	2.27 (1.91)	1.04 (1.42)	-0.253 (1.48)
ASC, transit[1]	2.00 (2.13)	-0.829 (1.60)	-4.36 (1.69)**
Alternative-specific variables			
Travel time by car (logged)	-0.253 (0.284)	-0.411 (0.542)	-0.697 (0.528)
Travel time by transit (logged)	-1.19 (0.354)***	-0.0931 (0.486)	-0.099 (0.516)
Travel time by active modes (logged)	-1.56 (0.252)***	-1.01 (0.457)**	-1.59 (0.517)***
Cost of commuting by transit	0.0121 (0.0625)	-0.144 (0.129)	0.0293 (0.127)
Socio-demographics			
Age (logged)	0.0108 (0.530)	-0.00389 (0.0144)	0.0252 (0.0158)
Number of cars owned by household	-0.136 (0.138)	-0.0271 (0.128)	0.0905 (0.144)
College education (yes/no)	-0.496 (0.207)**	-0.321 (0.196)	-0.505 (0.215)**
Errands during commute (yes/no)	-0.132 (0.264)	0.0531 (0.246)	-0.066 (0.277)
Household size	0.240 (0.0747)***	0.17 (0.069)**	0.159 (0.0785)*
Owns home	0.0757 (0.226)	-0.422 (1.30)	-0.768 (0.239)***
Income[2]	-0.409 (0.205)**	-0.377 (0.19)**	-0.00861 (0.219)
Zhengzhou local	0.0173 (0.204)	-0.501 (0.195)**	0.293 (1.36)
Male (yes/no)	-0.203 (0.202)	0.758 (1.31)	-0.585 (1.42)
Interactions			
Homeownership * travel time by car (logged)	--	0.656 (0.529)	--
Homeownership * travel time by transit (logged)	--	0.367 (0.415)	--
Homeownership * travel time by active modes (logged)	--	0.561 (0.470)	--
Homeownership * travel cost by transit	--	0.342 (0.124)***	--
Male * travel time by car (logged)	--	-0.681 (0.520)	0.29 (0.627)
Male * travel time by transit (logged)	--	-0.837 (0.407)**	0.186 (0.511)

Male * travel time by active modes (logged)	--	-0.816 (0.460)*	0.571 (0.589)
Male * travel cost by transit	--	-0.140 (0.107)	0.244 (0.142)*
Zhengzhou local * travel time by car (logged)	--	--	-0.19 (0.631)
Zhengzhou local * travel time by transit (logged)	--	--	-0.0818 (0.507)
Zhengzhou local * travel time by active modes (logged)	--	--	-0.251 (0.574)
Zhengzhou local * travel cost by transit	--	--	-0.369 (0.133)***
Model fit			
Likelihood ratio test[3]	129.9***	105.5***	72.5***
Adjusted R^2	0.268	0.139	0.370
Akaike Information Criterion (AIC)	903	1061	776
Bayesian Information Criterion (BIC)	972	1165	881
Observations	561	561	561

*Notes: *significant at 10% level. ** significant at 5% level. *** significant at 1% level. Each cell gives the estimated coefficient and, in parenthesis, the robust standard error. [1]ASC is the alternative-specific constant. Driving a personal motorized vehicle is the reference category, so it has no ASC and sociodemographic variables were excluded from its utility function to avoid multicollinearity. [2]Income is encoded by whether the respondent falls within the top or bottom half of the sample income distribution. As illustrated in Table 2, our sample respondents' income brackets were mainly in the RMB 50,000 – 150,000 and RMB 150,000 – 300,000 categories, so it was difficult to include more granular income groupings in the model. [3]Likelihood ratio test against the market share is of the listed model against a base model with only ASCs.*

The treatment effect is the key variable of interest given the study's randomized controlled experiment design. We do, however, also provide some discussion of the control variables included in the MNL.

Treatment: The treatment effect of carbon emission information is significant across all three experimental stages after controlling for other relevant variables. In all stages, it increases the utility of transit or active modes above the reference case of commuting by car. This indicates that social externality information is useful for altering individuals' modal choice intentions, and that it has an enduring effect over the course of the experiment. The magnitude of the coefficient does, however, diminish from 1.32 to 0.590 between Stages 1 and 3 as respondents received information about air pollution exposure and its detrimental health impacts. Since the car group is the reference category and the exponentiation of MNL coefficients represents change in outcome odds, this means that, if all other variables were held constant, the treatment increases the odds of the average car commuter choosing a non-car mode by 274% during Stage 1. Yet it only increases these odds by 177% in Stage 2 and 80% in Stage 3.[9] This diminishing treatment effect is expected from the narrowing gap between the lines in Figure 2, and shows how countervailing health and pollution severity information dampens the treatment effect.

[9] Impacts on outcome odds are calculated by exponentiating the relevant treatment coefficient. For Stage 1, $\beta_{treatment\_stage1}$ = 1.32, so the change in odds is 100*($e^{1.32}$-1) = 274%. This means holding all else constant, an individual receiving the social externality treatment in Stage 1 is 2.74 times more likely to try transit or active travel versus the reference case of car travel. Similarly, for Stage 2, 100*($e^{1.02}$-1) = 177% and for Stage 3, 100*($e^{0.59}$-1) = 80%.

Alternative-specific characteristics: Prior to Stage 3, when respondents were informed of how cars may protect riders against pollution exposure, the ASCs for transit and active modes were not significant. In Stage 3, transit's significant, negative ASC indicates that, unlike in previous stages, respondents now have an average preference for commuting by private car rather than by transit.

Next, we see in all three stages that the transit cost coefficients are insignificant. This is not surprising in the context of our specific study since we target car commuters, a subpopulation that may be, on average, less sensitive to the low transit cost typical in Zhengzhou when making stated choices about how to commute during the various stages of our experiment. The average commuting cost by transit in our sample is approximately 2 RMB, while the average petrol price is around 6 RMB per liter in Zhengzhou. As such, the average transit fare in our sample might be marginal enough to be arguably neglected for car users when making stated choices regarding mode shift. Car commuters in China also tend to be higher income, further suggesting lower sensitivity to small transit fares (He and Thogersen 2017, Zhao and Zhao 2020). In addition to the low fare costs, Zhengzhou's relatively flat fare system gives little variability in transit costs for us to capture significant sensitivity to price—95% of the transit costs are below 6 RMB, the IQR runs from 1 RMB to 3 RMB, and the standard deviation is 1.99 RMB. The insignificant coefficient of transit cost reflects these characteristics of our sample and study setting.

In all three stages, longer travel time for active modes significantly detracts from an individual's likelihood of choosing active travel rather than cars. Car travel time was not significant. Transit travel time was only significant for Stage 1. Car and transit travel times are typically significant in choice model results, yet our study differs from the typical choice modelling study in that it sampled only those who typically commute by car. They may be more accustomed to time spent commuting inside motorized vehicles like buses and cars. This is especially likely for their baseline mode of car commuting and along the routes that they regularly drive on weekdays. For commuting trips, prior research based on revealed preference data suggests that past experience with a specific transportation mode does play a significant role in determining the "experienced choice set" (the set of transportation modes used over a long period of time, e.g., over the course of six months in Ton et al., 2020) and in influencing mode choice behavior (Hensher and Ho, 2016). Separately, since our sample is not extremely large, our models may not have enough statistical power to detect smaller effects of car or transit travel time on the outcome once severe pollution and health impact information were given in Stages 2 and 3. This is especially a concern given that travel time variability for these modes are more limited— when logged to correct for a long right tail, the standard deviation for car travel time is 0.53 and for transit is 0.45, much smaller than 0.72 for active travel.

Sociodemographics: Sociodemographic factors related to household income or size are fairly consistently predictive of stated mode choice among this sample of drivers. College graduates are less likely to reduce their intention to drive in any stage, though the magnitude of the effect is smaller in Stage 2. Higher income households are also less likely to give up driving in the first two stages, although by Stage 3, this income effect is no longer significant. Homeowners are less likely to cut intention to commute by car in the last stage. These findings align with expectations; Chinese society strongly associates cars with high social status, so it is less likely that those of higher socioeconomic status would consider giving up the usage and public display of their cars. Meanwhile, results also indicate that in all stages, individuals from larger households are slightly more likely to switch to active and transit modes. This could be because individuals within larger, often multigenerational families have been conditioned by their home environments to be more attentive to societal needs. Lastly, being Zhengzhou locals only significantly lowered intention to switch away from cars in Stage 2. Age and running errands during a commute were not significant, potentially because the sample was dominated by younger workers who do not regularly include errands in their commutes.

The number of cars owned by the household was not a distinguishing factor at any stage. Only a very small number of respondents had no car—six in the control group and 13 in the treatment group. Therefore, the non-significance of the car ownership variable indicates that having *more* cars does not appear to impact modal intent; it does not necessarily capture the potentially larger effect of a household moving from zero cars to one car. The small number of non-car owners in the sample is also relevant to mitigating concerns surrounding one potential weakness of our study, which is that our response variable (and the related survey question) did not differentiate between those who regularly commuted by private car, versus by taxi or ridehailing. The small number of non-car households suggests the car commuters were most likely to drive rather than hire a ride, and that the changes in modal intention observed are not primarily caused by taxi and ridehail commuters without the fixed investment of car ownership. As detailed in footnote 8, this is confirmed by other survey questions showing the majority of our sample rarely commuted by hired cars.

Interaction terms between travel time/cost and sociodemographics: Interaction terms did not improve model fit for Stage 1, but did for Stages 2 and 3. In Stage 2, the interaction of gender with transit travel time, and the interaction of gender with active mode travel time, were both significant and negative with similar magnitudes. This suggests that once they were informed about the elevated local pollution level in Stage 2, men were more likely than other genders to stay with commuting by car when their travel time for undertaking the same commute by transit or walking/biking lengthened. This additional effect is above and beyond the individual effects of travel times and gender. Separately, the interaction of homeownership with transit travel cost was significant and positive. This indicates that, above and beyond the individual effect of transit cost or homeownership, a situation of rising transit costs does less to blunt homeowners' openness to switching to commuting by transit, than for those who do not own their homes. However, the extent of this impact in our data is limited, since transit costs only showed limited variability, with 95% of fares falling under 6 RMB and the full range covering 0-13 RMB.

In Stage 3, the interaction of transit cost with gender is significant and positive, suggesting that men are less deterred than other genders from trying to commute by transit when facing transit cost increases at this stage (i.e. men appear less sensitive to incremental transit costs). On the other hand, travel cost's interaction with being a Zhengzhou local was significant and negative, suggesting locals become less likely to switch from car to transit as transit cost increases (i.e. locals appear more sensitive to incremental transit costs).

5.4 Difference-in-difference results

Results from the individual-level difference-in-difference (DID) regression offer insight on how the rate of change in car commuting intention diverged between control and treatment group members from stage to stage. As explained in the Methods section, our response variable is a binary indicating whether the respondent still intended to commute by car at a particular experimental stage. A separate regression was run for each stage and the key coefficients of interest are the DID effects.

Table 4 Difference-in-difference regression results, intention to commute by car

Variable	1. Stage 1 β (robust S.E.)	2. Stage 2 β (robust S.E.)	3. Stage 3 β (robust S.E.)
Intercept	0.501 (0.220)**	-0.049 (0.308)	0.490 (0.305)
Experimental variables			
DID effect (for Stage 1, Stage 2, *or* Stage 3)	-0.259 (0.038)***	0.061 (0.055)	0.103 (0.054)*
Treatment group (yes/no)	0.001 (0.007)	-0.263 (0.038)***	-0.203 (0.041)***
Time effect (for Stage 1, Stage 2, *or* Stage 3)	-0.229 (0.024)***	-0.161 (0.037)***	0.193 (0.037)***
Alternative-specific variables			
Travel time by car (logged)	0.043 (0.036)	0.058 (0.052)	0.003 (0.051)
Travel time by transit (logged)	0.103 (0.041)**	0.140 (0.057)**	0.035 (0.056)
Travel time by active modes (logged)	-0.018 (0.034)	-0.031 (0.049)	0.002 (0.048)
Cost of commuting by transit	0.003 (0.005)	0.002 (0.007)	-0.004 (0.008)
Socio-demographics			
Age (logged)	0.003 (0.055)	0.034 (0.076)	-0.033 (0.008)
Number of cars owned by household	0.012 (0.014)	0.019 (0.020)	-0.003 (0.019)
College education (yes/no)	0.048 (0.021)**	0.089 (0.031)***	0.082 (0.030)***
Errands during commute (yes/no)	0.015 (0.027)	0.008 (0.038)	-0.006 (0.037)
Household size	-0.023 (0.007)***	-0.040 (0.011)***	-0.029 (0.010)***
Owns home	-0.005 (0.022)	0.035 (0.032)	0.103 (0.032)***
Income[2]	0.040 (0.019)*	0.078 (0.029)***	0.039 (0.029)
Zhengzhou local	-0.005 (0.020)	0.043 (0.029)	0.060 (0.028)**
Male (yes/no)	0.027 (0.020)	0.086 (0.029)***	0.059 (0.028)**
Model fit			
F-statistic	31.1***	10.35***	9.81***
Adjusted R^2	0.31	0.13	0.12
Observations (individual-stage)[1]	1122	1122	1122

*Notes: *significant at 10% level. ** significant at 5% level. *** significant at 1% level. Each cell gives the estimated coefficient and, in parenthesis, the robust standard error. [1]Number of observations is 1,122 for each regression because we have 561 respondents, and each of the three regressions compares two adjacent stages.*

Regression results affirm our earlier descriptive analysis. The significant Stage 1 time effect of -0.229 in Table 4 indicates individuals from *both* the control and treatment groups expressed, on average, significant decreases of 22.9 pp in intended car usage at Stage 1. The result reflects Anable and Gatersleben (2005) and Lofgren and Nordblom (2006) which found that simply giving car-based commuters information on alternative modes, as Stage 1 did for both groups, can encourage green travel adoption. Both groups experienced another, much smaller decrease in intended car use at Stage 2. The Stage 2 time effect is significant at -16.1 pp. This could stem from car commuters' recognition that cutting pollution is important in polluted areas like Zhengzhou (Garcia-Sierra 2015); or a desire to improve air quality for their own long-term health. It could also reflect social desirability bias, by which respondents wished to appear more pro-social in light of their city's polluted condition—but often reversed course when the survey later showed the personal health toll from pollution exposure (Grimm 2010). In Stage 3, the time effect reverses sign, becoming +19.3 pp. This shows that car usage intention rose for both groups once respondents were faced with the health impact of pollution exposure.

The treatment group variable is also significant with a coefficient of around -20 pp in the second and third regressions. This captures any general effect of being in the treatment group that holds constant from stage to stage (e.g., from Stage 1 to Stage 2 in regression 2). It reflects the observation stated earlier that across all stages beyond baseline Stage 0, the treatment group reported lower likelihood to commute by car.

We now come to the key results in the regressions—the DID effects. Unlike the time effects and the treatment variable, DID effects indicate how changes in car use intention differed between the control and treatment, *and* how this evolved over time. The Stage 1 DID coefficient (-0.259) is significant and shows that the treatment group made an additional 25.9 pp moved away from car usage beyond what was seen for the control group in Stage 1. In Stage 2, however, the insignificant DID coefficient indicates treatment had no additional impact beyond the control. For Stage 3, the DID coefficient is positive and significant, suggesting those who received the social externality treatment actually rebounded more steeply towards car use than did the control group by 10.3 pp. This suggests that green emission information interventions can have unexpected effects within the policy portfolio that weaken their net greening impact. That said, the cumulative DID effect summed across all stages is still negative and significant at -0.096, indicating there is still a net greening effect of the social externality treatment.

Among control variables, college education is positive and significant while household size is negative and significant for all three stages. Transit time is positive and statistically significant for the first two stages, while income, gender, homeownership and being a local are also occasionally significant. These coefficients again indicate heterogeneity in the sample that affects the degree to which car use intentions change when various pollution and public health information interventions are given. Future work with a larger sample could examine this heterogeneity with greater robustness.[10]

Combined with MNL analysis, DID suggests that while the social externality treatment lowers the overall level of intended car commutes, the degree to which the treatment outperforms the control group varies widely depending on the other pollution and health information affecting the respondent. This is a word of caution for governments managing multiple green transportation and public health information campaigns, as interactions between them can drastically diminish the effectiveness of an information campaign.

---

[10] In Appendix B.3, we also examined interaction effects to assess how differences in the sociodemographic and commute time characteristics of the sample may affect the changes individuals experienced in intention to commute by car across experimental stages. For Stages 1 and 2, the stage time effect interacted with either an individual's public transit, car, or active mode travel time was significant (we only present the results using the interaction with transit time). The addition of interaction effects did not significantly alter the DID coefficients, and interactions with the DID or treatment variables were not significant. This suggests that commuting times affect changes in how both control and treatment group members' intention to commute by car across the first two experimental stages. However, there is no additional interaction effect for the treatment group only.

6. Discussion and Conclusions

6.1 Key takeaways

This study suggests that giving car commuters information regarding the emissions social externality of their commute choice can encourage them to begin assessing non-car travel modes. The effectiveness of such information interventions in promoting intention to switch away from car use aligns with earlier empirical studies linking the communication of an act's emissions consequences to individuals' intention to use greener modes (Daziano, et al. 2017, Geng, Long and Chen 2016, Mir, et al. 2016, Waygood and Avineri 2018, Waygood and Avineri 2016).

However, this study also shows that the benefits and therefore the cost effectiveness of such greening information interventions may diminish when the government's policy portfolio also includes public health measures educating residents about local pollution, its health impacts, and methods to reduce pollution exposure. Previous research of air pollution prevention behavior in China has shown that Chinese residents are willing to take a wide range of actions to reduce their pollution exposure, ranging from small adaptations like mask-wearing to large adaptations like seeking housing in cleaner areas (Barwick, et al. 2019). Changing commuting habits is also one of these preventive measures, and drivers with full knowledge of the emissions consequences of their commute mode would still naturally be drawn towards the perceived safety of vehicular travel.

Amplifying the concern is our finding that the treatment group may rebound toward car travel more steeply than the control group. This suggests that the benefits gained through investing in social externality information interventions could deteriorate even more rapidly where governments feel more compelled to educate residents on the health consequences of pollution, and methods for preventing exposure.

Despite the rebounds in car use, the social externality treatment group still finished Stage 3 with a 10-percentage point lower likelihood of intending to commute by car than the control. This is positive news for policymakers, suggesting some ability for social externality information to persist through countervailing information. However, the most effective way forward may require policymakers to take integrated action like transit development and land use planning, to provide lifestyle options that slash emissions while reducing the pollution exposure and health risk of using low-emission modes.

6.2 Limitations and directions for future work

As a stated preference study, the behavioral experiment in this research is open to hypothetical bias, which is commonly documented across this type of work (Hensher 2010). Because respondents do not need to execute their stated intentions and may wish to appear socially responsible before experimenters, they may bias their response towards lower car usage, perceiving this to be the experimenters' intent. The use of a control group and DID help limit these effects, allowing us to eliminate bias that affected the control and treatment group to the same degree. However, they fail to capture factors that affect one group more than the other (e.g., the treatment group may become more susceptible to social approval bias than the control, after receiving social externality information). A follow-up experiment using randomized controlled trials and recording actual behavioral changes following information campaigns would address this issue.

Further, the study scenario is limited in the sense that it focuses on interactions between informational interventions. In reality, policymakers undertake a variety of measures, including transit subsidies or investments in enabling hard infrastructure like rail extensions.

Transit pass subsidies in particular are common interventions, and the interaction between this and information interventions are worth further exploration. Subsidies may, for example, crowd out the psychological greening benefit derived from green information. In general, synergy between multiple soft and hard measures, informational versus financial policies, remain ripe for investigation.

From a modeling perspective, the inclusion of estimated travel costs by car (rather than only by taxi) in the initial survey would also have been valuable to enable more thorough choice modelling. Relatedly, the inclusion of validated attitudinal constructs would have been useful for enabling latent class models and allowing us to investigate the underlying attitudes that may be tied to the intended mode choice decisions we observed. This could have helped us also to better capture systematic heterogeneity across respondent behavior.

The generalizability of this paper is also limited to young- and middle-aged urban workers due to the nature of our sample. Table 2 shows that the mean age of both the treatment and the control groups was 32, and the sample in general underrepresents those over the age of 45. Further, by sampling from companies in the CBD, this study focuses on workers rather than the general population. Future studies can take a broader sampling approach to check whether the results here extend to other populations and to non-commuting trips.[11]

The survey underlying this paper also did not track users to see how the social externality information intervention's impact evolved over time in reality. Erev and Barron (2005) found that people's evaluation of transportation modes changes with experience, as people tend to remember their most recent trip or the worst travel time trips, and rely on these memories more than official descriptive information such as transit time from Google Maps. Further research can track how mode change instigated by green information interventions evolve.

The presentation of pollution information in this paper used a neutral framing, and did not explore how prospect theory principles of asymmetry in loss versus gain framing could affect the effectiveness of the information shown, unlike Waygood and Avineri (2018), Mir *et al.* (2018) and others. Further study regarding psychological framing and driver attitudes may benefit policymakers from an efficacy and cost-effectiveness perspective. Additionally, the inclusion of rigorously measured attitudinal variables can allow us to better apply latent class models to capture heterogeneity and class membership among sample respondents, which can contribute to better nuance in policy decisions.

Related to the above, we note also that our final sample size is somewhat small with 561 respondents, since we filtered out non-car commuters from our survey experiment. A larger sample could improve the power of our estimates and support more complex models such as latent class models or mixed logit models to describe taste heterogeneity. Smaller samples are also more susceptible to bias; however, the lack of statistically significant differences between the control and treatment groups indicate that we were successful in randomizing participants into our controlled experiment. The use of a randomized controlled survey design helps to limit bias in our results.

This present study also does not evaluate the collective action dimension of the car use decision. A social dilemma lies at the core of the decision to refrain from car use—while driving is in the interest of the individual, it is detrimental to the collective interest through its environmental and congestion effects. People are more willing to stop driving if others stop, because only then would all parties realize the social benefit. Evaluations of how social context shapes an individual's reaction to social externality information would provide

[11] For example, non-commute trips are often more flexible than commute trips, giving travelers more ability to switch to other modes. In our sample's control and treatment groups, 85% of commuters reported having very little commute-time flexibility, and roughly 13% said they had only a little flexibility.

policymakers a sense of a “critical mass” that their policies would have to affect to engender widespread mode change (Lindenberg and Steg 2007).

6.3 Conclusion

Urban governments in polluted places often face a dual mandate to reduce emissions, while informing citizens about the hazards of local pollution and ways to limit personal pollution exposure. At times in the transportation sector, these mandates may come into conflict because the travel modes that are perceived to be the most protective against pollution, may not be the greenest options.

While our results suggest that policymakers can make headway simultaneously in both public health education and information campaigns supporting transit and active travel, it more importantly points to the necessity of taking broader measures to reduce air pollution, as other research has suggested from different angles (Jiang, et al. 2017, Li and Kamargianni 2017). Hard infrastructure improvements such as more frequent transit service or train line expansions can help residents translate intention to switch to transit into reality, by reducing travel time and cost. A more extensive transit system that reduces walking time would additionally cut commuter pollution exposure, as would mixed-use urban planning. In fact, Zhengzhou itself is moving in this direction, with the 2019 approval of seven subway extension projects (Xinhua News 2019). Further, enabling electric vehicles charged through clean energy can reduce the tradeoffs drivers perceive they face between self-protection from pollution and environmental stewardship. By paying close attention to the tradeoffs residents must make as they choose between greener and more polluting modes, and staying cognizant of interactions between programs in its policy portfolio, governments can make progress towards both protecting the health of urban residents and promoting green urban mobility.

ACKNOWLEDGMENTS

This project is supported by the MIT China Future City Zhengzhou Living Lab Program. Funding for data collection was provided by the MIT Sustainable Urbanization Lab and conducted in partnership with the municipal government of Zhengzhou, China.

REFERENCES

Ai, Chunrong and Edward C. Norton. 2003. “Interaction terms in logit and probit models.” *Economics Letters* 80: 123-129.

Ajzen, Icek. 1991. "The Theory of Planned Behavior." *Organizational Behavior and Human Decision Processes* 50: 179-211.

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Zou. 2019. "From Fog to Smog: the Value of Pollution Information." *NBER Working Paper No. 26541.*

Ben-Akiva, Moshe, and Steven R. Lerman. 1985. *Discrete Choice Analysis.* Cambridge, MA: MIT Press.

Brook, Robert, Sanjay Rajagopalan, C. Arden Pope III, Jeffrey Brook, Aruni Bhatnagar, Ana Diez-Roux, and Fernando Holguin. 2010. "Particulate Matter Air Pollution and Cardiovascular Disease." *Journal of the American Heart Association: Circulation* 2010 (121): 2331-2378.

Cepeda, Magda, Josje Schoufour, Rosanne Freak-Poli, Chantal Koolhaas, Klodian Dhana, Wichor Bramer, and Oscar Franco. 2017. "Levels of ambient air pollution according to mode of transport: a systematic review." *Lancet Public Health* 2: e23-34.

Chen, Shuai, Paulina Oliva, and Peng Zhang. 2017. "The Effect of Air Pollution on Migration: Evidence from China." *NBER Working Paper No. 24036.*

Chin, Michael. 2015. "Basic mechanisms for adverse cardiovascular events associated with air pollution." *BMJ Global Health: Heart* 101: 253-256.

Daziano, Ricardo, E.O.D. Waygood, Zachary Patterson, and Marketa Kolova. 2017. "Increasing the influence of CO2 emissions information on car purchase." *Journal of Cleaner Production* 164: 861-871. doi: 10.1016/j.jclepro.2017.07.001.

Dimick, Justin, and Andrew Ryan. 2014. "Methods for Evaluating Changes in Health Care Policy The Difference-in-Differences Approach." *Journal of the American Medical Association* 312 (22): 2401-2402.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. "New evidence on the impact of sustained exposure air pollution on life expectancy from China's Huai River Policy." *PNAS* (PNAS) 114 (39): 10384-10389. doi: 10.1073/pnas.1616784114.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, Peng Yin, and Maigeng Zhou. 2015. "Growth, Pollution, and Life Expectancy: China from 1991-2012." *American Economic Review* 105 (5): 226-231.

Gaker, David, and Joan Walker. 2013. "Revealing the Value of “Green” and the Small Group with a Big Heart in Transportation Mode Choice." *Sustainability* 5: 2913-2927. doi:10.3390/su5072913.

Garcia-Sierra, Marta et al. 2015. "Behavioural economics, travel behavior and environmental-transport policy." *Transportation Research Part D* 288-305.

Geng, Jichao, Ruyin Long, and Hong Chen. 2016. "Impact of information intervention on travel mode choice of urban residents with different goal frames: A controlled trial in Xuzhou, China." *Transportation Research Part A* 91: 134-147.

Gong, Mimi, Yin Shasha, Xingke Gu, Yuanqian Xu, Nan Jiang, and Ruiqin Zhang. 2017. "Refined 2013-based vehicle emission inventory and its spatial and temporal characteristics in Zhengzhou, China." *Science of the Total Environment* 599-600: 1149-1159.

Grimm, Pamela. 2010. "Social Desirability Bias." *Wiley International Encyclopedia of Marketing* doi: 10.1002/9781444316568.wiem02057.

Guo, Yuntao, and Srinivas Peeta. 2019. "Impacts of Personalized Accessibility Information on Residential Location Choice and Travel Behavior." *Travel Behaviour and Society* 19: 99-111.

Hamanaka, Robert, and Gokhan Mutlu. 2018. "Particulate Matter Air Pollution: Effects on the Cardiovascular System." *Frontiers in Endocrinology* 9 (680).

He, Sylvia, and John Thogersen. 2017. "The impact of attitudes and perceptions on travel mode choice and car ownership in a Chinese megacity: The case of Guangzhou." *Research in Transportation Economics* 62: 57-67.

Heinen, Eva. 2016. "Identity and travel behaviour: A cross-sectional study on commute mode choice and intention to change." *Transportation Research Part F: Traffic Psychology and Behaviour* 43: 238-253.

Hensher, David. 2008. "Climate change, enhanced greenhouse gas emissions and passenger transport – What can we do to make a difference?" *Transportation Research Part D: Transport and Environment* 2: 95-111.

Hensher, David. 2010. "Hypothetical bias, choice experiments, and willingness to pay." *Transportation Research Part B: Methodological* 44 (6): 735-752.

Hensher, David, Chinh Q. Ho. 2016. "Experience conditioning in commuter modal choice modelling—Does it make a difference?" *Transportation Research Part E* 95: 164-176.

Hensher, David, John M Rose, and William H. Greene. 2015. *Applied Choice Analysis.* Cambridge, U.K.: Cambridge University Press.

Jiang, Baoguo, Song Liang, Zhong-Ren Peng, Haozhe Cong, Morgan Levy, Qu Cheng, Tianbing Wang, and Justin Remais. 2017. "Transport and public health in China: the road to a healthy future." *Lancet* 390: 1781-1791.

Karaca-Mandic, Pinar, Norton, Edward C., and Bryan Dowd. 2011. "Interaction Terms in Nonlinear Models." *Health Services Research* 47(1): 255-274.

Kishimoto, Paul, Valerie Karplus, Min Zhong, Eri Saikawa, Xu Zhang, and Xiliang Zhang. 2017. "The impact of coordinated policies on air pollution emissions from road transportation in China." *Transportation Research Part D* 54: 30-49. https://doi.org/10.1016/j.trd.2017.02.012.

Klockner, Christian, and Thomas Friedrichsmeier. 2011. "A multi-level approach to travel mode choice – How person characteristics and situation specific aspects determine car use in a student sample." *Transportation Research Part F: Traffic Psychology and Behaviour* 14 (4): 261-277.

Lechner, Michael. 2010. "The estimation of causal effects by difference-in-difference methods." *Foundations and Trends in Econometrics* 4(3): 165-224.

Lelieveld, Jos, John Evans, M. Fnais, Despina Giannadaki, and Andrea Pozzer. 2015. "The contribution of outdoor air pollution sources to premature mortality on a global scale." *Nature* 525: 367-371.

Li, Weibo, and Maria Kamargianni. 2017. "Air Pollution and Seasonality Effects on Mode Choice in China." *Transportation Research Record Journal of the Transportation Research Board* 2634: 101-109.

Lindenberg, Siegwart, and Linda Steg. 2007. "Normative, Gain and Hedonic Goal Frames Guiding Environmental Behavior." *Journal of Social Issues* 63 (1): doi: 0.1111/j.1540-4560.2007.00499.x.

Löfgren, A., and K. Nordblom. 2006. *The importance of habit formation for environmental taxation. Working Papers in Economics No. 204.* Department of Economics, Göteborg University.

Mir, Hussein, Behrang Koorosh, Mohammad Isaai, and Pegah Nejat. 2016. "The impact of outcome framing and psychological distance of air pollution consequences on transportation mode choice." *Transportation Research Part D* (46): 328-338.

Muller, Richard, and Elizabeth Muller. 2015. *Air Pollution and Cigarette Equivalence.* http://berkeleyearth.org/archive/air-pollution-and-cigarette-equivalence/.

Nawrot, Tim, Laura Perez, Nino Kunzli, Elke Munters, and Benoit Nemery. 2011. "Public health importance of triggers of myocardial infarction: a comparative risk assessment." *Lancet* 377 (9767): 732-40.

Prochascka, J.O., C.A. Redding, and K.E. Evers. 2008. "The transtheoretical model and stages of change." In *Health behavior and health education: Theory, research, and practice*, by K. Glanz, B.K. Rimer and Viswanath K., 97-121. Jossey-Bass.

Puhani, Patrick. 2012. "The treatment effect, the cross difference, and the interaction term in non-linear 'difference-in-differences' models." *Economics Letters* 115: 85-87.

Raaschou-Nielsen, Ole, Zorana Andersen, Rob Beelen, Evangelia Samoli, Massimo Stafoggia, Gudrun Weinmayr, Barbara Hoffmann, et al. 2013. "Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE)." *The Lancet: Oncology* 14 (9): 813-822.

Saberian, Soodeh, Anthony Heyes, and Nicholas Rivers. 2017. "Alerts work! Air quality warnings and cycling." *Resource and Energy Economics* 49: 165-185.

Song, Congbo, Jianjun He, Lin Wu, Taosheng Jin, Xi Chen, Ruipeng Li, Peipei Ren, Li Zhang, and Hongjun Mao. 2017. "Health burden attributable to ambient PM 2.5 in China." *Environmental Pollution* 223: 575-586.

Sun, Cong, Matthew E. Kahn, and Siqi Zheng. (2017). “Self-protection investment exacerbates air pollution exposure inequality in urban China.” *Ecological Economics: The Journal of the International Society for Ecological Economics* 131: 468–474.

Tainio, Marko, Audrey Nazelle, Thomas Gotschi, Sonja Kahlmeier, David Rojas-Rueda, Mark Nieuwenhuijsen, Thiago Henrick de Sa, Paul Kelly, and James Woodcock. 2016. "Can air pollution negate the health benefits of cycling and walking?" *Preventive Medicine* 87: 233-236.

Ton, Danique, Shlomo Bekhor, Oded Cats, Dorine Cuives, Sascha Hoogendoorn-Lanser, Serge Hoogendoorn. 2020. “The experienced mode choice set and its determinants: Commuting trips in the Netherlands.” *Transportation Research Part A* 132: 744-758.

Tu, Meng, Bing Zhang, Jianhua Xu, and Fangwen Lu. 2020. "Mass media, information and demand for environmental quality: Evidence from the “Under the Dome”." *Journal of Development Economics* 143.

Wang, Jian, Jianan Huang, Linlin Wang, Cuicui Chen, Dong Yang, Meiling Jin, Chunxue Bai, and Yuanlin Song. 2017. "Urban particulate matter triggers lung inflammation via the ROS-MAPK-NF-κB signaling pathway." *Journal of Thoracic Disease* 9 (11): 4398-4412.

Waygood, E., and E. Avineri. 2016. "Communicating transportation carbon dioxide emissions information: Does gender impact behavioral response?" *Transportation Research Part D: Transport and Environment* 48: 187-202.

Waygood, E., and Erel Avineri. 2018. "CO2 valuence framing: Is it really any different from just giving the amounts?" *Transportation Research Part D* (63): 718-732.

Woodcock, James, Phil Edwards, Cathryn Tonne, Ben Armstrong, Olu Ashiru, David Banister, Sean Beevers, and Zaid Chalabi. 2009. "Public health benefits of strategies to reduce greenhouse-gas emissions: urban land transport." *The Lancet* 374 (9705): 1930-1943.

Xinhua News. 2019. "China approves 7 subway projects in Zhengzhou." China Daily, April 28.

Yang, Ming, and John Zacharias. 2016. "Potential for revival of the bicycle in Beijing." *International Journal of Sustainable Transportation* 10 (6): 517-527.

Zhao, Zhan, and Jinhua Zhao. 2020. "Car pride and its behavioral implications: an exploration in Shanghai." *Transportation* 47: 793-810.

Zivin, Joshua Graff, and Matthew Neidell. 2018. "Air pollution's hidden impacts." *Science* 359: 39-40.

APPENDIX A: EXPERIMENTAL DESIGN INFORMATION

A.1 Sampling methodology

We used a stratified and clustered sampling methodology, with the industries based in Zhengzhou as the strata. We randomly selected companies from each stratum, such that the industry breakdown in our final sample of 93 local firms was approximately representative of their shares in the Zhengzhou economy according to census data. All employees in each of the selected firms were eligible for the survey (i.e., we clustered on the firms). Employees who were willing to participate were pre-screened to ensure they commuted regularly by car.

The actual survey took roughly 15-20 minutes to complete (though we originally budgeted 20- to 30-minute sessions) and was executed on an iPad using the Qualtrics interface for the questionnaire, alongside our customized website for the presentation of the information interventions. Survey staff consisted of 60 local graduate students, who were on-site to aid the distribution of the iPads and to monitor the survey area. They were trained to remain neutral in their interactions with the survey respondents, and to quietly note the survey ID numbers of respondents who appeared to rush through or be distracted during the survey. These survey responses were assumed to be invalid and deleted from the final responses in real-time. The entire survey was presented in Mandarin Chinese, the local language.

A.2 Response variable measurement

Modal intention was collected for Stages 1-3 using a question based on this template:

Given the information you have seen, what method would you choose for traveling to and from work next month?

- Personal motorized vehicle: My own car/hired car
- Public transit: Bus/subway
- Electric bike
- Bike: shared
- Bike: personal
- Walking
- Other (please specify)

The Stage 0 pre-screening question collecting baseline commuting mode was formatted similarly, but the question was worded as “How do you usually commute to and from work?” The survey was administered in Mandarin Chinese (the above question is an English translation). The idea of asking for traveler modal intentions for an upcoming commute trip is drawn from Klockner and Friedrichsmeier (2011). In addition to asking directly at each stage for the respondent’s commute mode choice, we also asked whether their decisions would change if offered a green subsidy of an unspecified amount for transit or active modes. If so, we asked what subsidy amount they would need to make the switch. Because the wording of these two latter questions (available on GitHub) did not produce highly reliable measures for willingness-to-accept analysis, we excluded them from this paper.

A.3 Pollution exposure calculations

The pollution factor and inhalation factor table below was built based on 1) the research team’s field monitoring of pollution levels along major commute routes to the Zhengzhou CBD, and 2) literature on inhalation rates.

Table A.3.1 Pollution factors by mode from field monitoring in Zhengzhou, inhalation factors by mode from literature

	Pollution factor (field monitoring)	Inhalation factor (literature)
Car (with AC)	0.8	0.16
Bike	1	1
Walk	1	1
Bus	1.1	0.72
Subway	1.2	0.49

From this, the exposure of individual *i* using transportation mode *j* for one commute trip was calculated as $Exposure_{ij} = AmbientPM2.5 \times PollutionFactor_j \times InhalationFactor_j \times CommuteTime_{ij}$

A.4 Polluted day illustration

In Stage 2 of the survey, respondents were presented public health information about the hazardous level of air pollution in Zhengzhou, prior to being asked about their intended commute mode next month. The screenshot from Qualtrics below shows the photo of a polluted day in Zhengzhou that was given to respondents.

APPENDIX B: ADDITIONAL DIFFERENCE-IN-DIFFERENCE RESULTS

We also conducted a preliminary DID analysis without controls, following the method in Geng et al (2016). Because control and treatment group membership in our experiment was randomized, the DID values calculated here should be close to those in the DID regressions with a full set of controls. Comparing results Table B.2 below against Table 4, we see this is indeed the case. For this preliminary analysis, $\bar{Y}_0^T$ and $\bar{Y}_1^T$ are the average outcome values in the treatment group for the previous and current stages (i.e., pre- and post- information intervention) respectively, while $\bar{Y}_0^C$ and $\bar{Y}_1^C$ are the average outcomes in the control group for those stages (Dimick and Ryan 2014). The DID estimator is thus $(\bar{Y}_1^T - \bar{Y}_0^T) - (\bar{Y}_1^C - \bar{Y}_0^C)$ and reflects the difference in slopes between the stages in Figure 2. We use t-tests to assess whether the DID effect is significant, and thus whether significant differences exist in how the treatment versus control group changed their modal intentions between experimental stages.

Table B.1: Methodology for preliminary DID estimation at each stage

	Previous stage (Pre)	Current stage (Post)	Post-Pre Difference
Treatment	$\bar{Y}_0^T$	$\bar{Y}_1^T$	$\bar{Y}_1^T - \bar{Y}_0^T$
Control	$\bar{Y}_0^C$	$\bar{Y}_1^C$	$\bar{Y}_1^C - \bar{Y}_0^C$
Treatment-Control	$\bar{Y}_0^T - \bar{Y}_0^C$	$\bar{Y}_1^T - \bar{Y}_1^C$	$(\bar{Y}_1^T - \bar{Y}_0^T) - (\bar{Y}_1^C - \bar{Y}_0^C)$

The results of interest are column 7: DID. Comparing this column to the DID coefficient estimates in the regression with controls in Table 4 shows that our DID estimates are fairly robust across these specifications, confirming successful randomization in our experiment.

Table B.2 Preliminary difference-in-difference on intended mode use, using t-tests

		Control group (n = 280)			Treatment group (n = 281)			DiD
Mode	**Stage**	**Before (1)**	**After (2)**	**Difference (3): (2)-(1)**	**Before (4)**	**After (5)**	**Difference (6): (5)-(4)**	**(7): (6) - (3)**
Car	Stage 1	1.000	0.771	-0.229***	1.000	0.512	-0.488***	-0.259***
	Stage 2	0.771	0.611	-0.161***	0.512	0.413	-0.0996***	0.061
	Stage 3	0.611	0.804	0.193***	0.413	0.708	0.295***	0.103***
Transit	Stage 1	0.00	0.154	0.154***	0.00	0.285	0.285***	0.131***
	Stage 2	0.154	0.289	0.136***	0.285	0.381	0.096***	-0.04
	Stage 3	0.289	0.129	-0.161***	0.381	0.189	-0.192***	-0.031
Active	Stage 1	0.00	0.075	0.075***	0.00	0.203	0.203***	0.128***
	Stage 2	0.075	0.100	0.025	0.203	0.206	0.004	-0.021
	Stage 3	0.100	0.068	-0.032	0.206	0.103	-0.103***	-0.071***

*Note: *** significant at 1% level*

Table B.3 DID regression with interaction effects, intention to commute by car

Variable	1. Stage 1 β (robust S.E.)	2. Stage 2 β (robust S.E.)	3. Stage 3 β (robust S.E.)
Intercept	0.967 (0.199)**	-0.352 (0.323)	0.362 (0.333)
Experimental variables			
DID (Stage 1, Stage 2, *or* Stage 3)	-0.266 (0.038)***	0.065 (0.055)	0.104 (0.054)*
Treatment group (yes/no)	0.004 (0.005)	-0.265 (0.038)***	-0.204 (0.041)***
Time effect (Stage 1, Stage 2, *or* Stage 3)	-1.16*** (0.164)	0.448* (0.244)	0.447 (0.251)*
Alternative-specific variables			
Travel time by car (logged)	0.043	0.058	0.003

	(0.036)	(0.052)	(0.051)
Travel time by transit (logged)	-0.014 (0.034)	0.215 (0.062)***	0.068 (0.066)
Travel time by active modes (logged)	-0.018 (0.033)	-0.030 (0.049)	0.002 (0.048)
Cost of commuting by transit	0.003 (0.004)	0.002 (0.007)	-0.004 (0.008)
Socio-demographics			
Age (logged)	0.003 (0.054)	0.034 (0.075)	-0.033 (0.075)
Number of cars owned by household	0.012 (0.013)	0.019 (0.020)	-0.003 (0.019)
College education (yes/no)	0.048 (0.021)**	0.087 (0.030)***	0.082 (0.030)***
Errands during commute (yes/no)	0.015 (0.027)	0.008 (0.038)	-0.006 (0.037)
Household size	-0.023 (0.007)***	-0.040 (0.011)***	-0.029 (0.010)***
Owns home	-0.005 (0.022)	0.035 (0.032)	0.103 (0.032)***
Income[2]	0.040 (0.019)**	0.078 (0.029)***	0.039 (0.029)
Zhengzhou local	-0.005 (0.020)	0.043 (0.029)	0.060 (0.028)**
Male (yes/no)	0.027 (0.020)	0.085 (0.029)***	0.059 (0.028)**
Interaction			
Travel time by transit (logged) * Time effect (Stage 1, Stage 2, *or* Stage 3)	0.233 (0.040)***	-0.152 (0.060)**	-0.064 (0.062)
Model fit			
F-statistic	31.8***	10.14***	9.30***
Adjusted R^2	0.32	0.12	0.13
Observations (individual-stage)[1]	1122	1122	1122

*Notes: *significant at 10% level. ** significant at 5% level. *** significant at 1% level. Each cell gives the estimated coefficient and, in parenthesis, the robust standard error. [1]Number of observations equals 1,122 for each regression because we have 561 respondents, and each of the three regressions above compares two adjacent stages (e.g. regression 2 above compares Stage 2 results to Stage 1 results to assess the change between these two time periods).*