The New York congestion charge

—

Pre-analysis plan*

Stefano Carattini†‡§ Jonathan Hall¶ Axel Ockenfels‖
Klaus Schmidt** Ravi Vora††

December 10, 2024

*The original design of this study dates of 2023, with the details being finalized ahead of the scheduled launch of the New York congestion charge in June 2024. The pause of the congestion charge meant that we paid for scripting charges but could hold off the survey administration. The project was revived following Governor Hochul's decision in November 2024 to unpause the congestion charge. The project's code name on OSF is Beliefs and Information in Gotham – Assessing the Popularity of Pricing through Learning via Experience (BIG APPLE).

†Department of Economics, Georgia State University, corresponding author: scarattini@gsu.edu.
‡CEPR
§CESifo
¶Culverhouse College of Business, University of Alabama
‖Department of Economics, University of Cologne
**Department of Economics, Ludwig Maximilian University of Munich
††Department of Economics, University of Wyoming

Abstract

This project aims at combining policy evaluation of the New York congestion charge with the causal analysis of public support. The congestion charge in New York provides a unique opportunity to examine people's understanding of and support for congestion charges both before and after the implementation of the policy, compared to a control group. That is, it allows at the same time learning about the policy (policy evaluation) and learning about how voters learn about the policy and potentially revise their beliefs accordingly (causal analysis of public support). Policymakers are likely to care about both items and to be more likely to experiment with ex-ante unpopular policies in presence of not only more evidence about their effectiveness, but also about their potential ex-post popularity. The New York congestion charge also provides the ideal framework to examine the role of information provision as a substitute, and complement, to experience.

1 Introduction

Many policies that experts recommend are not implemented by policymakers, because of their (ex-ante) popularity. Congestion charges are a case in point. This project aims at combining policy evaluation and the causal analysis of public support (Carattini et al. 2024) applied to the New York congestion charge and inform policymakers about two key items: how the congestion charge works (policy evaluation) and how voters learn about how it works and, as a result, possibly reconsider their stance vis-à-vis the policy (causal analysis of public support).

With policy evaluation, our objective is to contribute to a relatively limited strand of research that examines the impact of congestion charges on key outcomes such as congestion, accidents, local air pollution, and housing values (Gibson and Carnovale 2015; Green et al. 2016, 2020; Tang and van Ommeren 2020; Tang 2021; see also Leape 2006 for a descriptive analysis applied to London). In line with the existing literature, we use administrative data for policy evaluation purposes. However, we also use survey data, which give us much more granular information on a variety of possible margins of adjustment in commuting behavior and impacts on travel satisfaction following the implementation of the congestion charge. Hence, our objective is to provide extensive evidence on the functioning of the New York congestion charge, a policy that affects the largest metropolitan area in the United States and one of the largest cities in the world.

With the causal analysis of public support, our objective is to provide evidence on how beliefs about and public support for congestion charges may evolve with direct experience of the policy, thus contributing to existing causal evidence on belief revision in the field (Carattini et al. 2018), anecdotal evidence from before-after studies (Schuitema et al. 2010; Andersson and Nässén 2016), and causal evidence from the lab (Cherry et al. 2014; Dal Bó et al. 2018; Janusch et al. 2021).

This project also tests the extent to which information may be a substitute for experience. Following evidence showing belief revision with experience, several studies on environmental taxes have provided information to respondents, as a potential substitute for experience, on how such taxes would work, aiming at addressing potential biased beliefs (i.e. overly pessimistic beliefs) about their economic, environmental, or distributional effects. This strand of research, starting with Carattini et al. (2017) and including, among others, Dechezleprêtre et al. (2022), shows promise for the potential of informational treatments to contribute to address belief revision, including across countries and cultures. An informational treatment was also used in Dal Bó et al. (2018) and in research specific to congestion charges, which we summarize in Section 5. Informational treatments have been used also to try to address biased beliefs about other policies or social issues (see Haaland et al. 2023 for a non-exhaustive review).

However, information may also be a complement to experience, to the extent that the effects of a policy are not entirely salient to voters (with respect to the counterfactual). Hence, we are also interested in the extent to which information (provided at baseline) may be a complement to experience.

To test the role of substitutability or complementarity of information with respect to experience, the baseline survey includes a randomized informational treatment, which describes the experience of frontrunner cities with congestion charges.

The survey covers respondents in the New York metropolitan area, respondents in four other metropolitan areas in the United States which currently do not have a congestion charge and have no immediate plans to introduce one, as well as in Greater London and in Singapore, which both already have a congestion charge and so where voters' beliefs about congestion charges may be less affected by New York's experience with congestion charges. That is, the survey combines untreated control metropolitan areas and already-treated control metropolitan areas.

In short, the experimental design is as follows. Respondents are surveyed twice through two survey waves, to form a survey panel. At baseline, half of the respondents across metropolitan areas are exposed to the informational treatment. Between waves, respondents in New York's metropolitan area are exposed to the New York congestion charge. Hence, part of the sample is exposed to information only, part of the sample to experience only, part of the sample to both information and experience, and part of the sample to neither.

The remainder of this document is organized as follows. Section 2 provides details about the local context and economic rationale for our study, in particular concerning the causal analysis of public support and the informational treatment. Section 4 describes the data that this project uses. Section 5 outlines the empirical approaches employed in the project.

2 Background

2.1 Local context

Congestion and local air pollution from road traffic are among the most important issues in many cities around the world. Driving and congestion create an important externality, i.e. an external cost to society, due to extended journeys, noise, accidents, and local air pollution, coupled also with greenhouse gas emissions (Small et al. 2005; Small and Verhoef 2007; Li 2012; Jacobsen 2013). The United Nations (2018) predict huge increases in the absolute number and fraction of the world population living in cities, in developed and developing countries alike. According to the OECD (2017), total motorized mobility may increase by about 40% (90%) by 2030 (2050) with respect to 2015. Whether higher urban population translates into higher traffic congestion depends on public policy. Congestion is a textbook example of a negative externality. Pigou (1920) already called for corrective taxes, now known as "Pigouvian," to address negative externalities. In the 1960s, several seminal studies provided guidelines on how to precisely internalize the externality of driving: pricing road traffic, for instance through congestion charges.

Economists often complain that while the adoption of carbon pricing is steadily picking up (World Bank 2024), adoption is still too slow to tackle the climate externality. However, still according to the World Bank (World Bank 2024), some 70 jurisdictions around the world price carbon at the moment. Unlike the climate externality, traffic congestion is a very local issue, with immediate harms. Yet only a handful of cities in the world have congestion charges. That is, there are way fewer cities that have a congestion charge than jurisdictions pricing carbon. Not much has changed since the 1960s, when William S. Vickrey made the following statement: "in no other major area are pricing practices so irrational, so out of date, and so conducive to waste as in urban transportation" (Vickrey 1963, p. 452).

Lukewarm public support is likely the main reason for the very limited adoption of congestion charges around the world (Gu et al. 2018). Congestion charges were rejected in the United Kingdom, in Birmingham, Edinburgh, and Manchester. They were also rejected in the Swedish city of Gothenburg, although in a non-binding referendum. Voters in another Swedish city, Stockholm, voted in favor of the congestion charge, but only after a trial period, which we know substantially reduced local air pollution (Simeonova et al. 2019). Policymakers in New York also considered congestion charges in the past, before giving up. Early plans date from the first decades of the 20^{th} century, while a 2006 proposal by former mayor Bloomberg was eventually abandoned at the state level, for lack of public support. More recently, however, the proposal was revived. After another failed attempt at introducing legislation implementing a congestion charge in 2015, a new proposal dating of 2017 from former state governor Cuomo evolved over time until it was passed by the state legislature, as part of the state budget, in 2019. The proposal implied that the entire island of Manhattan south of 60^{th} street would be charged, with a few notable exceptions circumventing the city center: the Franklin D. Roosevelt East River Drive, the West Side Highway, and the Battery Park Underpass. The policy was to be implemented as a cordon charge, with cameras at each entry to the inner zone (but not inside the zone capturing movements within the island). Vehicles were expected to be tolled once per day only. Exemptions applied for emergency vehicles and vehicles transporting passengers with disabilities. The charge level was expected to be between $9 and $23 during peak hours, higher if not paying through E-ZPass. Revenues were expected to be used mostly to fund public transportation. Tax deductions for low- and moderate-income households were also foreseen. The congestion charge was expected to be launched on June 30, 2024, before Governor Hochul called it off in early June 2024. Governor Hochul did not abandon the policy completely, but froze it for an undetermined period.

In November 2024, the policy was unpaused. The policy design remained largely unchanged except for one major difference. The congestion charge is now being introduced with a price of $9 for entering the zone. The congestion charge is expected to reach gradually the previous target of $15 towards the end of the decade. Prices were adjusted downward also for ride-hailing companies. The reference price is for drivers using the electronic toll system E-ZPass. 110 toll readers are present in the city for processing the congestion charge. The congestion charge is expected to enter into force on January 5, 2025.

2.2 Economic rationale

The project builds on the following premises. Limited information and biased beliefs among citizens are among the reasons for the insufficient adoption of ambitious environmental policies. While the early literature on public support for environmental taxes did not really question people's responses, even concluding, as done in Steg, Dreijerink, and Abrahamse (2006), that subsidies encouraging the adoption of specific behaviors would be more effective than environmental taxes because people perceive them as such, the second generation of studies started highlighting important information asymmetries between economists and citizens. Knowledge gaps and biased beliefs became evident when scholars started looking at how experience of a given Pigouvian policy changed people's perceptions, and in turn public support. Studies initially provided anecdotal evidence, for instance from the Swedish cities of Stockholm (Schuitema et al. 2010) and Gothenburg (Andersson and Nässén 2016), which both implemented congestion charges after a trial period. More systematic analyses followed shortly after, providing what is now called the causal analysis of public support (Carattini, Dur, and List 2024). Using data for 2012 and 2013, Carattini et al. (2018) built a panel to measure how perceptions about and public support for pricing garbage by the bag may vary before and after the policy's implementation, compared to an already-treated control group, while also providing a policy evaluation exercise. Support increases substantially with experience, while no difference is observed in the already-treated control group, which was already familiar with the policy. To the best of our knowledge, Carattini et al. (2018) is still the only study combining policy evaluation with a systematic analysis of public support. Experience has also been shown to increase public support for Pigouvian taxes in lab experiments (Cherry et al. 2014; Dal Bó et al. 2018; Janusch et al. 2021).

If experience can lead to belief revision, it may follow naturally that trial periods should be used more often, as advocated in Carattini et al. (2018). However, trial periods also need some degree of public support. Further, they may work better for policies whose effects are very salient and where the short-term and long-term price elasticities of demand do not vary substantially. Hence, a natural extension to this literature was to test whether information provision could be a substitute, even if imperfect, for experience. Carattini et al. (2017) analyzed with survey data the drivers of and barriers to public support for a carbon tax submitted to Swiss citizens in 2015. Misperceptions played an important role in its rejection. Public support would have been higher if revenues were earmarked for environmental purposes, because this is how many people think an environmental tax can be effective. At the same time of the ballot, a discrete choice experiment was administered to another sample of voters. Voters were informed about economic, distributional, and environmental effects of different carbon tax designs, simulated with a computable general equilibrium model. Since then, many additional studies have provided information to survey respondents and lab participants, in a randomized fashion. However, we still do not know to what extent information and experience may be substitutes and (or) complements. Hence, the proposed research aims to shed substantial new light on belief revision following real-world experience, advancing the frontier in the literature.

This research project examines the implementation of the New York congestion charge, measuring belief revision about the policy in New York versus a group of control cities. It also measures behavioral change, thus combining analysis of public support with policy evaluation. As mentioned, the only existing study combining policy evaluation and analysis of public support is Carattini et al. (2018) and uses data from 2012 and 2013. Our research question and policy of interest are different, the sample size is larger, and the set of outcomes broader.

3 Design

The core of this study is a representative survey panel, aimed at examining beliefs and public support but also conceived for policy evaluation purposes. Respondents are recruited in the metropolitan statistical area (MSA) to which New York City belongs, including individuals living inside and outside the area under the congestion charge, as well as in the other top metropolitan statistical areas in the country (Boston–Cambridge–Newton, Chicago–Naperville–Elgin, Philadelphia–Camden–Wilmington, and Washington–Arlington–Alexandria), serving as an untreated control group, and in Greater London, United Kingdom, and the city-state of Singapore, both providing an already-treated control group. We do not expect beliefs about congestion charges to change in London or Singapore following the implementation of the congestion charge in New York, but they may vary in the United States, a feature that this design would allow us to document.

This design leverages the fact that respondents in and around New York City will directly experience the congestion charge, with the other MSAs serving as control units. The survey panel is obtained with two survey waves, about six months apart. Each wave has largely the same survey instrument, to measure variation within subjects on the same outcomes. At the same time, we are interested in the role of information.

Hence, at baseline, half of the respondents in each MSA, including New York City, are exposed to an informational treatment about congestion charges, a feature that allows us to examine how good of a substitute and (or) complement information is to experience. To examine substitutability, we compare baseline beliefs under information provision in control cities and ex-post beliefs in New York City. To examine complementarity, we analyze ex-post beliefs in New York City for the respondents subject to the informational treatment against (i) baseline (and ex-post) beliefs in control cities, for respondents exposed to the information treatment (information only), and (ii) ex-post beliefs in New York City for respondents not exposed to the informational treatment (experience only).

The informational treatment to which half of the respondents across geographies are exposed details the effects of congestion charge policies implemented in the past and the costs of congestion. This informational treatment will be displayed to treated subjects prior to eliciting their support for congestion charges. The intervention reads as follows (based on Green et al. 2016; Green et al. 2020, Traffic Mobility Review Board 2023; Transport for London 2023):

> *A congestion charge is a fee charged on cars and motor vehicles being driven within central parts of a city during busy hours. After implementing congestion charges, London, Milan, Singapore, and Stockholm each experienced 20-30% reductions in road traffic. As a result, driving time required to reach a destination in the central business district fell substantially, air quality improved, and the number of accidents fell. Many commuters shifted from driving to the use of public transportation. In large cities, excess congestion is estimated to cost businesses, commuters, and residents substantial amounts of money. For example, in New York excess congestion is estimated to cost $20 billion a year.*

The control group is instead exposed to the following text:

> *A congestion charge is a fee charged on cars and motor vehicles being driven within central parts of a city during busy hours.*

Further, given some expected attrition, which the marketing company can minimize but cannot completely avoid, a pool of new respondents is added during the second wave, to detect potential experimenter demand effects in the panel. Absent experimenter demand effects, ex-post answers in the panel should be comparable with ex-post answers in the new pool of respondents. Section 4 discusses about sample selection, sample size, and attrition. Section 5 presents our main empirical specifications.

Policy evaluation is conducted using survey data as well, comparing respondents in and around New York City with control MSAs. Our design and sample size are such that potential influence, if any, between the informational treatment and survey responses about commuting patterns would not interfere with the policy evaluation exercise, as robustness tests can be run with respondents not exposed to information. Further, policy evaluation using survey data is complemented with administrative data (traffic counts, pollution from monitoring stations and satellite data, severe traffic accidents). Survey data can augment administrative data by measuring behavioral change along the intensive and extensive margins as well as changes in utility (for instance due to reduced comfort from commuting via public transportation) that are not observed in administrative data. Once more, Section 4 discusses about sample selection, sample size, and attrition. Section 5 presents our main empirical specifications.

The survey aims to be representative of all the MSAs that are sampled, as well as of London and Singapore for the respective samples. The survey measures a wide variety of outcomes, including understanding, perceptions, narratives, and mental models related with congestion charges as well as many items related with commuting behaviors for policy evaluation purposes. The following sections provide more details.

4 Data

Our data will primarily consist of surveys conducted in December 2024, before the policy introduction estimated around January 5, 2025, and in June of 2025, about six months after the policy introduction. Six months is a period of time that gives respondents in New York City enough time to experience the effects of the policy and is still within the time frame suggested by survey companies for a panel study (attrition becoming an increasing concern beyond six months).[1]

[1] Given some expected attrition, the marketing company company will obtain new respondents in the second wave of the survey in order to balance sample sizes across the waves. Having a mix of new respondents and original respondents will have the added benefit of allowing us to test for

The samples for the initial survey will include 4,000 respondents in the New York-Newark-Jersey City metropolitan statistical area (MSA), the region that is being treated with the congestion policy introduction; 1,125 respondents each in the Boston, Chicago, Philadelphia, and Washington, D.C., MSAs, which are control MSAs that have never been treated with a congestion charge; 1,500 and 2,000 respondents each in Singapore and Greater London, respectively, which are control cities that are already treated in the past with congestion policy introductions. Sample sizes are generally determined by the size of the panel of potential respondents available to the survey company with which we collaborate for a given location, with such availability being a key criterion in the choice of the survey company.

Respondents will be compensated a fixed amount for participation and will be made aware of strong financial incentives to respond to the second survey wave in mid-2025. The second survey will consist of both respondents who participated in the first wave, as well as new respondents, to maintain sample size constant. Sample sizes and proportions will be approximately identical between waves. Respondents will be recruited by marketing research firm Bilendi, which will also host the survey and collect response data. Bilendi is instructed to recruit respondents with the goal of achieving sample representativeness for each metropolitan area on population measures of age and gender. Attrition rates are expected around 50%, with some variation across markets.

To maintain a desired quality level, respondents will be automatically removed from the study for completing the survey in too short a time to have properly read the questions, or for failing to pass simple attention checks that are included in the survey instrument. As detailed in the survey instrument, respondents will be asked standard demographic questions; open response questions concerning opinions about traffic and congestion; questions concerning their current commuting and transportation habits (including detailed trip information); their satisfaction with their transportation habits and options; open response questions concerning their knowledge about congestion charges and narratives about how they work; support for such congestion charges and their perception of the effects on them (including forecasts of the impact of the policy on their own choices) as well as on their metropolitan area (including distributional and competitiveness effects).

experimenter demand effects that could result from surveying repeat respondents twice. We can test whether the responses of the newly added respondents is meaningfully different from the responses of repeat respondents.

While respondents encounter all the questions, certain details of the question are tailored to the policies and conditions unique to the various geographies included in the sample. A few follow-up questions depend on the answer provided in the previous questions.

Power calculations for hypothesis tests are provided in Appendix Section B. The power calculations also describe our assumptions concerning attrition in the panel. In short, we expect some attrition but aim to minimize it. The degree of attrition eventually determines the sample size for the panel. Each survey wave, however, is conceived to have the same number of respondents, implying fresh respondents replacing those who may attrite. As mentioned, the inclusion of fresh respondents allows us to test for experimenter demand effects, i.e. whether respondents exposed to the New York congestion charge and asked about it respond differently than respondents exposed to the New York congestion charge but interviewed only ex post.

To collect the data necessary for the causal analysis of public support, we collect responses to our surveys on the following outcomes:

- Binary support for the congestion charge policy and degree of confidence about that support or opposition.
- Preferred level, in local currency, of the congestion charge.
- Second-order beliefs about preferences by eliciting beliefs about the percentage of one's metropolitan area that would support the policy.

We also collect data on preferred uses of the revenue generated from the congestion charge, the preferred policy level conditional on the preferred use of revenues, and degree of (perceived) knowledge about the policy. In addition to these general measures of policy support, we elicit a Likert rating concerning the policy's fairness and the specific policy impact forecast by the respondents on their own household's welfare and travel usage, as well as aggregate effects on commuting times, air pollution, economic activity, welfare of of all people and of low-income people in one's city and in one's metropolitan area. All of these responses will be used as outcome variables to provide rich detail about the causal impact on support and the mechanisms underlying it.

For policy evaluation purposes, we plan to combine data from our (panel) survey and administrative data from other sources (including other surveys). We aim to evaluate the impact of the congestion pricing zone on outcomes related to travel speeds, travel behavior, locational choices, (severe) accidents, and local air pollution. Specific outcomes we will measure include:

- Travel speeds and their inverse (time per kilometer) by time of day and for different geographies (inside the central business district – congestion charge zone in New York City, right outside the central business district, further from the city), and response times for emergency vehicles
- Travel volumes (car, transit, ride-hail)
- Changes in where people travel (inside versus outside the central business district) for same trip purpose
- Changes in why people travel (trip purpose)
- Changes in when people travel
- Changes in how people travel, including ride hail
- Traffic accidents and their severity
- Air quality (local air pollution)
- Locational choices
- Travel satisfaction

The following administrative data sources can complement our proprietary survey data: transit ridership data for the treated city and control cities;[2] ride-hail ridership data for New York City and Chicago;[3] station-specific data and subway origin-destination ridership estimates;[4] data on travel time collected from Google Maps for

[2] These data are available from the National Transit Database, https://www.transit.dot.gov/ntd (last accessed, December 6, 2024).

[3] These data are available at https://www.nyc.gov/site/tlc/about/tlc-trip-record-data.page and https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips-2023-/n26f-ihde/about_data (last accessed, December 6, 2024).

[4] These data are available from the Metropolitan Transportation Authority at https://data.ny.gov/Transportation/MTA-Subway-Hourly-Ridership-Beginning-July-2020/wujg-7c2s/about_data and https://data.ny.gov/Transportation/MTA-Subway-Origin-Destination-Ridership-Estimate-2/jsu2-fbtj/about_data (last accessed, December 6, 2024).

all trips as well as emergency vehicles response times;[5] traffic volume data as provided by the Highway Performance Monitoring System[6] as well as by the cities of interest themselves, including for pedestrian and bicycle traffic;[7] air quality data (air quality index as well as specific pollutants associated with traffic) from the monitoring stations as well as from satellite data; migration data from the American Community Survey and Current Population Survey.

5 Empirical approach

5.1 Causal analysis of public support

Our study focuses on two main channels through which people's understanding of and support for congestion charges can change: information and experience. We experimentally vary the exposure to information about congestion charges at baseline and measure the impact of that information immediately (first wave) and its potential persistence over time (second wave). We are unable to experimentally vary exposure to congestion charges (experience). To evaluate the role of experience, we use a quasi-experimental approach, comparing between the pre- and post-surveys how support and beliefs evolve in New York City, which is treated with a congestion charge, and the six control cities that have either never been treated or had already been treated with the implementation of a congestion charge in the distant past.

Because our study allows for two types of treatment, information and experience, we are also able to test the interactions of these, giving us a view into whether and

[5] Data from Google Maps can be complemented with additional data provided by the cities of interest themselves. For instance, New York posts travel time data at https://data.cityofnewyork.us/Transportation/DOT-Traffic-Speeds-NBE/i4gi-tjb9/about_data (last accessed, December 6, 2024). Emergency vehicles response times are available from https://www.usfa.fema.gov/nfirs/access-data/ (last accessed, December 6, 2024) as used in Brent and Beland (2020).

[6] These data are available from https://www.fhwa.dot.gov/policyinformation/hpms.cfm (last accessed, December 6, 2024)

[7] For instance, New York provides automated traffic volume counts at https://data.cityofnewyork.us/Transportation/Automated-Traffic-Volume-Counts/7ym2-wayt/about_data (last accessed, December 6, 2024), pedestrian counts at https://data.cityofnewyork.us/Transportation/Bi-Annual-Pedestrian-Counts/2de2-6x2h (last accessed, December 6, 2024), and bicycle counts at https://data.cityofnewyork.us/Transportation/Bicycle-Counts/uczf-rk3c/about_data (last accessed, December 6, 2024).

when information and experience are complements or substitutes:

$$Y_{ict} = \alpha_i + \theta_t + \beta_1\ information_i + \beta_2\ experience_{ct} + \beta_3\ information_i * experience_{ct} + \varepsilon_{ict} \tag{1}$$

where Y_{ict} is a binary variable coding the expressed policy support of individual i in metropolitan area c at survey wave t, with 1 corresponding to support and 0 to opposition. The variable α_i is an individual fixed effect, while θ_t is a survey wave fixed effect that is coded as 0 for the survey wave occurring in December 2024 and as 1 for the survey wave occurring in early 2025. The variable $information_i$ is the randomly assigned information treatment variable, where 1 corresponds to having received the information treatment described in Section 4 at baseline and 0 corresponds to having received the control condition. The variable $experience_{ct}$ is the nonrandom binary treatment variable, which is coded as 1 for all responses from London and Singapore and for responses from the second wave of the New York metropolitan area. All others are coded as 0. Given that the outcome variable is binary, we would test the robustness of our linear probability model result by also including a probit specification.

We can perform a similar analysis using our continuous measure of public support for a congestion charge, the preferred policy level (i.e. the fee). Because this response is provided in local currency, we must first normalize the responses to median income levels in the respective countries. After doing so Y_{ict}, the country-normalized preferred congestion fee of individual i in city c at survey wave t can be treated as a continuous variable in the regression specification of equation 1. Other continuous measures relevant for public support include our elicitation of second-order beliefs measuring the perceived policy's popularity. The outcome variable measuring second-order beliefs, i.e. a respondent's beliefs about the policy support among others, is one of the relevant measures that we plan to study. Other more specific measures concerning beliefs about the policy's impacts can be assessed in a similar way. As noted in Section 4, we collect a Likert rating concerning the policy's fairness and perceived policy impact forecast by the respondents on their own household's welfare and travel usage, as well as aggregate effects on commuting times, air pollution, economic activity, welfare of of all people and of low-income people in one's city and in one's metropolitan area. All of these responses will be used as the Y_{ict} outcome variable to provide rich detail about the causal impact on support and the mechanisms underlying it. We address multiple hypothesis testing in Section 5.3.

Many of the secondary outcome variables of interest mentioned above could be considered mediators of primary policy support measures, such as the binary policy support response and the preferred policy level. To consider these secondary variables' role as mediators, we can analyze the results as two-stage experiments. If we are considering the effect of the policy's introduction and want to better understand the mechanism from the secondary response variables, we can estimate the following two equations:

$$Y_{ict} = \alpha_i + \theta_t + \gamma_{dy}\ experience_{ct} + \gamma_{my}\ mediator_i + \varepsilon_{ict} \tag{2}$$

$$mediator_{ict} = \eta_i + \iota_t + \gamma_{dm}\ experience_{ct} + \nu_{ict} \tag{3}$$

Considering these two specifications, we can estimate the direct effect of experience on the primary support variable with γ_{dy} and the indirect effect of experience, channeled through the mediating secondary variable, as $\gamma_{my}\gamma_{dm}$.

Previous studies have measured the impact of information treatments on support for congestion policies. Baranzini et al. (2021) find that information about congestion pricing's impact on congestion and on pollution increased support to 53% and 57%, respectively, from a control of 51%, in Geneva. The difference in public support in the control group and in the treatment around pollution is very small at low charge rates, when public support is relatively high, but increases with policy stringency, becoming statistically significant. Arlinghaus et al. (2024) find that informational videos separately focusing on the policy's impacts on pollution, climate change, and time savings improved support by 11.4%, 7.1%, and 6.5%, respectively. These results were measured across two cities, Berlin and Paris, and were strongly statistically significant. We refer to these studies to compare our estimates concerning the impact of the informational treatment at baseline with priors.

5.2 Policy evaluation

In this section, we describe the main empirical approaches for the policy evaluation component of our study. Recall that this component uses both proprietary survey data and administrative data (whose original source may be another survey). The main approach that we plan to use with the survey data – and by extension also with the administrative data – consists of a canonical difference-in-difference model where metropolitan areas in the United States provide an untreated (and never-treated) control group to the New York metropolitan area, which itself starts out as untreated but becomes treated with the implementation of the congestion charge. While the inclusion of London and Singapore as already-treated control metropolitan areas is motivated mostly by the causal analysis of public support, we still plan to have specifications considering these metropolitan areas as control units as well. In the policy evaluation exercise, individual respondents from the survey panel – or individual records from administrative data such as a subway station's traffic – represent an observation, allowing for the use of individual-specific fixed effects as standard with canonical difference-in-differences. Standard errors are to be clustered at the metropolitan-area level, which is the level at which treatment is assigned. Although with less power, robustness tests include specifications that omit observations of respondents exposed to the informational treatment as a conservative approach in case such exposure would affect their behavior.

Since only one metropolitan area receives treatment, we also consider specifications suited for contexts with a single treated cluster, in particular synthetic control (Abadie and Gardeazabal 2003; Abadie et al. 2010, see also Abadie 2021 for a review) and synthetic difference in differences (Arkhangelsky et al. 2021). These approaches are especially suitable for policy evaluation with administrative data, where long time series exist. Metropolitan areas beyond those included in the survey can be part of the donor pool. Synthetic difference in differences also allows for the inclusion of already-treated control metropolitan areas, such as London and Singapore but also Milan, Gothenburg, and Stockholm, in case levels differ, accompanied by a discussion about dynamic effects.

For outcomes for which data for control cities cannot be accessed (for instance because not collected), we use an event study design leveraging the data for New York City.

5.3 Multiple hypothesis testing

The specifications for causal analysis of public support involve two distinct treatments and their effects on multiple outcomes of interest detailed in Sections 4 and 5. Multiple outcomes are also considered in the policy evaluation exercise. To address the potential of type-I errors that arise from multiple hypotheses, we will perform the Holm-Bonferroni Stepdown Correction (Holm, 1979) on the regression estimands of interest, denoted by β's in Equation 1. This iterative process involves collecting all of the p-values across specifications and to begin applying the Bonferroni criterion to the lowest p-value, comparing its value to the ratio of the significance threshold α divided by the total number of hypothesis tests, n. If we reject the null for the first test, we proceed to the following p-value, comparing it with $\frac{\alpha}{n-1}$. If we reject the null, the next p-value is compared with $\frac{\alpha}{n-2}$, and so on. This iterative process continues until we reach a p-value for which we fail to reject the null hypothesis. Any remaining p-values would be considered statistically insignificant under this correction.

References

Abadie, Alberto (2021) "Using synthetic controls: Feasibility, data requirements, and methodological aspects," *Journal of Economic Literature*, 59 (2), 391–425.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010) "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American Statistical Association*, 105 (490), 493–505.

Abadie, Alberto and Javier Gardeazabal (2003) "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, 93 (1), 113–132.

Andersson, David and Jonas Nässén (2016) "The Gothenburg congestion charge scheme: A pre–post analysis of commuting behavior and travel satisfaction," *Journal of Transport Geography*, 52, 82–89.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager (2021) "Synthetic difference-in-differences," *American Economic Review*, 111 (12), 4088–4118.

Arlinghaus, Johanna Brigitte, Théo Konc, Linus Mattauch, and Stephan Sommer (2024) "The effect of information framing on policy support: Experimental evidence from urban policies."

Baranzini, Andrea, Stefano Carattini, and Linda Tesauro (2021) "Designing effective and acceptable road pricing schemes: evidence from the Geneva congestion charge," *Environmental and Resource Economics*, 79 (3), 417–482.

Brent, Daniel and Louis-Philippe Beland (2020) "Traffic congestion, transportation policies, and the performance of first responders," *Journal of Environmental Economics and Management*, 103, 102339.

Carattini, Stefano, Andrea Baranzini, and Rafael Lalive (2018) "Is taxing waste a waste of time? Evidence from a supreme court decision," *Ecological Economics*, 148, 131–151.

Carattini, Stefano, Andrea Baranzini, Philippe Thalmann, Frédéric Varone, and Frank Vöhringer (2017) "Green taxes in a post-Paris world: Are millions of nays inevitable?" *Environmental and Resource Economics*, 68 (1), 97–128.

Carattini, Stefano, Robert Dur, and John List (2024) "Policy evaluation and the causal analysis of public support," *Science*, 386 (6721), 490–492.

Cherry, Todd L., Steffen Kallbekken, and Stephan Kroll (2014) "The impact of trial runs on the acceptability of environmental taxes: Experimental evidence," *Resource and Energy Economics*, 38 (C), 84–95.

Dal Bó, Ernesto, Pedro Dal Bó, and Erik Eyster (2018) "The demand for bad policy when voters underappreciate equilibrium effects," *Review of Economic Studies*, 85 (2), 964–998.

Dechezleprêtre, Antoine, Adrien Fabre, Tobias Kruse, Bluebery Planterose, Ana Sanchez Chico, and Stefanie Stantcheva (2022) "Fighting climate change: International attitudes toward climate policies," Working Paper 30265, National Bureau of Economic Research.

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

Green, Colin P., John S. Heywood, and María Navarro (2016) "Traffic accidents and the London congestion charge," *Journal of Public Economics*, 133, 11–22.

Green, Colin P, John S Heywood, and Maria Navarro Paniagua (2020) "Did the London congestion charge reduce pollution?" *Regional Science and Urban Economics*, 84, 103573.

Gu, Ziyuan, Zhiyuan Liu, Qixiu Cheng, and Meead Saberi (2018) "Congestion pricing practices and public acceptance: A review of evidence," *Case Studies on Transport Policy*, 6 (1), 94–101.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart (2023) "Designing information provision experiments," *Journal of Economic Literature*, 61 (1), 3–40.

Holm, Sture (1979) "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, 65–70.

Jacobsen, Mark R. (2013) "Fuel economy and safety: The influences of vehicle class and driver behavior," *American Economic Journal: Applied Economics*, 5 (3), 1–26.

Janusch, Nicholas, Stephan Kroll, Christopher Goemans, Todd L. Cherry, and Steffen Kallbekken (2021) "Learning to accept welfare-enhancing policies: an experimental investigation of congestion pricing," *Experimental Economics*, 24 (1).

Leape, Jonathan (2006) "The London Congestion Charge," *Journal of Economic Perspectives*, 20 (4).

Li, Shanjun (2012) "Traffic safety and vehicle choice: quantifying the effects of the 'arms race' on American roads," *Journal of Applied Econometrics*, 27 (1), 34–62.

OECD (2017) "ITF Transport Outlook 2017,"Technical report, OECD Publishing.

Pigou, A. C (1920) *The Economics of Welfare*, London: Macmillan and Co.

Schuitema, Geertje, Linda Steg, and Sonja Forward (2010) "Explaining differences in acceptability before and acceptance after the implementation of a congestion charge in Stockholm," *Transportation Research Part A: Policy and Practice*, 44 (2), 99–109.

Simeonova, Emilia, Janet Currie, Peter Nilsson, and Reed Walker (2019) "Congestion pricing, air pollution, and children's health," *Journal of Human Resources*, 0218.

Small, Kenneth A., Clifford Winston, and Jia Yan (2005) "Uncovering the distribution of motorists' preferences for travel time and reliability," *Econometrica*, 73 (4), 1367–1382.

Small, Kenneth and Erik Verhoef (2007) *The Economics of Urban Transportation*, New York: Routledge.

Tang, Cheng Keat (2021) "The cost of traffic: Evidence from the London congestion charge," *Journal of Urban Economics*, 121, 103302.

Tang, Cheng Keat and Jos van Ommeren (2020) "Accident externality of driving: Evidence from the London congestion charge," *Journal of Economic Geography*, 22 (3), 547–580.

Traffic Mobility Review Board (2023) "Recommendations for Congestion Pricing Toll Rate Schedule."

Transport for London (2023) "Congestion Charge Marks 20 Years of Keeping London Moving Sustainably," `https://`

tfl.gov.uk/info-for/media/press-releases/2023/february/congestion-charge-marks-20-years-of-keeping-london-moving-sustainably.

United Nations (2018) "World Urbanization Prospects: The 2018 Revision,"Technical report, United Nations.

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

World Bank (2024) "State and Trends of Carbon Pricing – 2024,"Technical report, World Bank.

Appendix

A Survey instrument

The survey instrument is made available separately on the same platform.

B Power calculations

B.1 Binary outcomes

This study considers both public support and policy evaluation. The primary binary public support variable of interest is support for the congestion charge. Examples of important binary policy evaluation include a decision to change one's departure time or use of public transportation in response to the congestion charge. With the goal of conservatism in our power calculations, we assume the variance maximizing outcome proportion to be 0.5.

To improve understanding of the measurements of interests, we provide the following descriptions, using the binary policy support variable as a guiding example: (i) the effect of an informational treatment on New York area residents on wave 1 support for a congestion charge; (ii) the effect of an informational treatment on wave 1 support for a congestion charge for the entire sample; (iii) the effect of an informational treatment on New York area residents on wave 2 support for a congestion charge; (iv) the effect of an informational treatment on wave 2 support for a congestion charge for the entire sample; (v) the effect of experience in the New York area of the policy between waves 1 and 2 on support for a congestion charge, using partly a panel of respondents who responded in both waves but received no informational treatment during wave 1; (vi) the effect of experience in the New York area of the policy between waves 1 and 2 on support for a congestion charge, considering only the subset that received an informational treatment; (vii) the difference in the effect of experience between informational treatment and control group; (viii) the difference in the effect of experience in the New York area between waves 1 and 2 and the effect of information on wave 1 support in the full sample.

Using 24,000 interviews over the 2 waves, with 8,000 interviews coming from New York area respondents and 16,000 interviews coming from outside the New York area, a wave 1 informational treatment probability across all respondents of 0.5, and an expected retention probability between waves 1 and 2 of 0.5, power calculations are performed using standard power level of 0.8 and significance level of 0.05, corresponding to 2.8 standard deviations given reasonably sized samples.

B.1.1 Case i

Sample: 4,000, with 2,000 treated.

Difference of means, assuming no other covariates

Baseline of binary outcome variable: $\sim 50\%$

Variance of individual outcomes:

$$\sigma^2 = \frac{p(1-p)}{n_0} + \frac{p(1-p)}{n_1}$$

For power level of 0.8 and significance level of 0.05, $Z_{\alpha/2} + Z_\beta = 2.8$

$$MDE = 2.8\sqrt{\frac{0.25}{2000} + \frac{0.25}{2000}} = 0.044$$

B.1.2 Case ii

The effect of information across the entire sample.

Sample: 12,000, with 6,000 treated.

$$MDE = 2.8\sqrt{\frac{0.25}{6000} + \frac{0.25}{6000}} = 0.026$$

B.1.3 Case iii

Assuming 50% retention rate between waves and 50% probability of assignment to the informational treatment, measuring the effect of the informational treatment on wave 2 responses of New York area respondents:

$$MDE = 2.8\sqrt{\frac{0.25}{1000} + \frac{0.25}{3000}} = 0.051$$

B.1.4 Case iv

Under similar assumptions, applying case iii for all respondents:

$$MDE = 2.8\sqrt{\frac{0.25}{3000} + \frac{0.25}{9000}} = 0.030$$

B.1.5 Case v

For the difference in differences on the effect of policy experience in the New York City area alone, and continuing to assume 50% retention rate between the wave 1 (prior to policy introduction) and wave 2 (following policy introduction):

$$MDE = 2.8\sqrt{0.5\left(\frac{0.25}{2000} + \frac{0.25}{3000}\right) + 0.5\left(\frac{0.25}{4000} + \frac{0.25}{6000}\right)} = 0.035$$

The calculation assumes that residual variance with the inclusion of individual fixed effects or with the inclusion of geographic/sociodemographic controls is 0.4 and 0.6 of the outcome variance, respectively. With a 50% retention rate between waves, this results in a residual-to-total variance ratio of 0.5.

B.1.6 Case vi

Performing the same calculation, but for the subsample randomized into the informational treatment:

$$MDE = 2.8\sqrt{0.5\left(\frac{0.25}{2000} + \frac{0.25}{1000}\right) + 0.5\left(\frac{0.25}{4000} + \frac{0.25}{2000}\right)} = 0.047$$

B.1.7 Case vii

For the difference in the difference-in-difference coefficient between the informational treated and information untreated subsamples:

$$MDE = 2.8\sqrt{0.5\left(\frac{0.25}{2000}+\frac{0.25}{1000}\right)+0.5\left(\frac{0.25}{4000}+\frac{0.25}{2000}\right)+0.5\left(\frac{0.25}{2000}+\frac{0.25}{1000}\right)+0.5\left(\frac{0.25}{4000}+\frac{0.25}{2000}\right)}$$

$$= 0.059$$

B.1.8 Case viii

Comparing experience alone to informational treatment alone:

$$MDE = 2.8\sqrt{1\left(\frac{0.25}{6000}+\frac{0.25}{6000}\right)+0.5\left(\frac{0.25}{2000}+\frac{0.25}{3000}\right)+0.5\left(\frac{0.25}{4000}+\frac{0.25}{6000}\right)}$$

$$= 0.043$$

B.2 Continuous outcomes

In this section, we extend the power calculations to continuous variables. An example for public support would be elicited response of the preferred level of the congestion charge, which we define as x below. We assume a standard deviation, or σ_x of \$10. Examples for policy evaluation would be the effect of the congestion charge on commute time outcomes or commute satisfaction outcomes. We use commute time outcomes, defined as y, to guide the examples below and use a standard deviation, σ_y of 30.8 minutes, as measured in the 2019 American Time Use Survey.

B.2.1 Case i

Using $\sigma_x = 10$,

$$MDE_x = 2.8\sigma_x\sqrt{\frac{1}{2000}+\frac{1}{2000}} = \$0.88$$

Thus we are powered to detect a \$0.88 effect on the preferred congestion charge level from our informational treatment, given \$10 standard deviation.

B.2.2 Case ii

$$MDE_x = 2.8\sigma_x\sqrt{\frac{1}{6000} + \frac{1}{6000}} = \$0.51$$

B.2.3 Case iii

$$MDE_x = 2.8\sigma_x\sqrt{\frac{1}{1000} + \frac{1}{3000}} = \$1.02$$

B.2.4 Case iv

$$MDE_x = 2.8\sigma_x\sqrt{\frac{1}{3000} + \frac{1}{9000}} = \$0.59$$

B.2.5 Case v

$$MDE_x = 2.8\sigma_x\sqrt{0.5\left(\frac{1}{2000} + \frac{1}{3000}\right) + 0.5\left(\frac{1}{4000} + \frac{1}{6000}\right)} = \$0.70$$

Thus we are powered to detect a \$0.70 effect on the preferred congestion charge level from experience treatment.

Case v is most relevant policy evaluation setting. Using $\sigma_y = 30.8$ min:

$$MDE_y = 2.8\sigma_y\sqrt{0.5\left(\frac{1}{2000} + \frac{1}{3000}\right) + 0.5\left(\frac{1}{4000} + \frac{1}{6000}\right)} = 2.16 \text{ min}$$

Thus we are powered to detect a 0.0 minute effect on commute times from the congestion charge policy. This is 5% of the mean of 43.5 minutes.

B.2.6 Case vi

$$MDE_x = 2.8\sigma_x \sqrt{0.5\left(\frac{1}{2000}+\frac{1}{1000}\right)+0.5\left(\frac{1}{4000}+\frac{1}{1000}\right)} = \$1.03$$

$$MDE_y = 2.8\sigma_y \sqrt{0.5\left(\frac{1}{2000}+\frac{1}{1000}\right)+0.5\left(\frac{1}{4000}+\frac{1}{1000}\right)} = 3.20 \text{ min}$$

B.2.7 Case vii

$$MDE_x = 2.8\sigma_x \sqrt{0.5\left(\frac{1}{2000}+\frac{1}{1000}\right)+0.5\left(\frac{1}{4000}+\frac{1}{2000}\right)+0.5\left(\frac{1}{2000}+\frac{1}{1000}\right)+0.5\left(\frac{1}{4000}+\frac{1}{2000}\right)}$$
$$= \$1.18$$

$$MDE_y = 2.8\sigma_y \sqrt{0.5\left(\frac{1}{2000}+\frac{1}{1000}\right)+0.5\left(\frac{1}{4000}+\frac{1}{2000}\right)+0.5\left(\frac{1}{2000}+\frac{1}{1000}\right)+0.5\left(\frac{1}{4000}+\frac{1}{2000}\right)}$$
$$= 3.63 \text{ min}$$

B.2.8 Case viii

$$MDE_x = 2.8\sigma_x \sqrt{1\left(\frac{1}{6000}+\frac{1}{6000}\right)+0.5\left(\frac{1}{2000}+\frac{1}{3000}\right)+0.5\left(\frac{1}{4000}+\frac{1}{6000}\right)}$$
$$= \$0.86$$