Welfare Estimates of Shifting Peak Travel*

Robert W. Hahn, Robert D. Metcalfe, Eddy H.F. Tam

August 23, 2023

Abstract

We develop novel estimates of peak and off-peak price elasticities for urban mass transit demand in San Francisco using a large natural experiment with 3.6 million trip sessions and a natural field experiment that both have exogenous price subsidies. We then estimate the welfare impacts for these price subsidies using a sufficient statistics approach. Our analysis suggests that off-peak subsidies can increase welfare, but the positive effects are reduced when consumers take the decisions of others into account compared to when they do not. We also find a large variation in the welfare impacts of shifting travel to different periods, which is explained by differences in demand and congestion characteristics. Finally, we show that the targeting of subsidies can increase welfare, but need not do so if the regulator does not have accurate information on demand.

Keywords: sufficient statistics, marginal value of public funds, welfare, congestion, regulation, benefit-cost analysis.

*Hahn: University of Oxford & Technology Policy Institute; Metcalfe: University of Southern California & NBER; Tam: King's College London. We would like to thank Ryan Greene-Roesel at Bay Area Rapid Transit (BART) and the BART team for partnering with us to conduct this research and taking a leading role developing the experiments reported in this paper. We also thank the Federal Highways Agency and the San Francisco County Transportation Authority (especially Joe Castiglione and Camille Guiriba), and BART for funding the program. We greatly appreciate the thoughtful comments of Leo Bursztyn, Thomas Chaney, Raj Chetty, Nathan Hendren, Jonathan Libgober, Victor Mylonas, Matthew Tarduno, Shosh Vasserman, Reed Walker, Zi Yang Kang, and seminar participants at Harvard University, University of California, Berkeley, and the University of Southern California. We also thank those at Metropia who helped to implement the field experiment, especially Yi-Chang Chiu and Ali Arian. The usual caveat applies.

1 Introduction

There are many problems in which economic welfare may be increased through shifting demand across peak and off-peak periods (Boiteux, 1960). Applications include automobile travel, mass transit, electricity, energy and telecommunications. In this paper, we focus on mass transit. The general problem we address is estimating the economic welfare impacts of a price change on consumption in the peak and off-peak periods. Our welfare estimates rely on well-identified demand elasticity estimates within and across time periods. We also allow for prices to differ from marginal cost and for externalities in consumption, such as congestion and pollution.

Using a large natural experiment and a natural field experiment, we develop novel estimates of peak and off-peak price elasticities for urban mass transit demand in San Francisco. The experiments were conducted by the Bay Area Rapid Transit (BART) in California.[1] Both experiments allow us to obtain causal estimates of the price elasticities of demand. The first, which includes 17,500 BART riders (7% of BART riders) and over 3.6 million trip sessions, is a natural experiment with a price change in the off-peak period for six months. Riders were staggered into the pricing program over the course of six weeks. Prices were reduced an hour before the peak period (630am-730am) and an hour after the peak period (830am-930am).

The second is a natural field experiment with randomized prices across 1,900 riders over four months (i.e., 77,000 trip sessions). The experiment incentivized treated customers to shift their riding demand into the least congested train within a 40 minute time window from their usual departure time by lowering prices in the new time window. With detailed consumer level data on demand and individual characteristics, both of these experiments yield own and cross price elasticities, which are used in the welfare analysis. Across both experiments, we have train-level data on the weight of the train and the time the consumer enters and leaves a station, allowing us to measure congestion. We estimate the welfare impacts of the price subsidies in the experiments using a sufficient statistics approach (Chetty, 2009).

The subsidies in both experiments influenced consumer choice in the expected direction. They reduced travel in the period that did not receive the subsidy, and increased travel in the period that received the subsidy. We estimated an own price elasticity for the period in which consumers received the subsidy and a cross-price elasticity for the period in which they did not. The own and cross price elasticities are around -0.86 and 0.44 respectively in the natural experiment and -0.94 and 0.54 respectively in the field experiment. These estimates suggest

[1] BART is a rail and subway system that serves the greater San Francisco area. It is the urban public transit system that Daniel McFadden used to estimate various discrete choice models that contributed to his Nobel prize work (McFadden, 1974; Domencich and McFadden, 1975; McFadden et al., 1977; McFadden, Tye and Train, 1977; McFadden, 2001).

some departure time flexibility for BART users. Other non-experimental elasticity estimates for the BART system as a whole are comparable to our own elasticity for off-peak demand (See, e.g., McFadden (1974)). But our own elasticity estimate is a little higher in absolute value than the -0.6 to -0.75 range in the review of estimates in the US by Litman (2004) and Holmgren (2007), and those estimated in Mexico by Davis (2021).[2]

Our two experiments offer different insights into the impact of providing subsidies at particular times for economic welfare. We find that the subsidy in the natural experiment generally increases welfare. In the base case, we show that the marginal value of public funds (MVPF) is 1.6 in trying to shift travel to the off-peak period from the peak period (730am-830am).[3] This means that for each additional dollar of net cost to the government, the marginal benefit to all parties is $1.60. We also calculate the net benefits per dollar of subsidy in the natural experiment, and find that the net benefits are $0.36 per dollar of subsidy expenditure.[4]

In the field experiment, we had price variation across more time periods and show that the welfare benefits varied dramatically across routes and time periods due to differences in demand and congestion characteristics. In particular, we show that net benefits per person in the field experiment are sometimes less than zero. We also show that some MVPFs in the field experiment are less than one, which means that the willingness to pay associated with the policy is less than the net cost to the government.[5]

We do several sensitivity analyses to examine the robustness of our main findings. In particular, we examine three sets of behavioral assumptions about consumer response to see how they would change our welfare results. The first relates to how welfare results would change if we scaled up the subsidy to all BART customers, but allowed for some customers to be less responsive to the subsidy than others, perhaps because of selection effects. The second relates to measuring the impact on consumer welfare under the assumption that consumers take the full impact of others' behavior into account in making their transit choices–a case we refer to as endogenous congestion.[6] In both the scaling and endogenous congestion simulations, the effect on per capita welfare is to reduce the net benefits from the subsidy by up to 47%. The third assumption relates to the level and shape of the congestion damage function. We find

[2]We have not found other well-identified *cross-price* elasticity estimates for peak and off-peak in public transit, and thus do not provide a comparison

[3]See Hendren (2020) and Finkelstein and Hendren (2020) for excellent reviews of this literature. Our MVPF of 1.6 is quite favorable in comparison to the marginal cost of public funds of around 0.8 to 1 through changes in the linear tax rate (Hendren, 2020).

[4]The net benefits measure we use in our welfare model is based on total net benefits for a discrete change in the subsidy, and is in dollars. In contrast, the MVPF is a marginal measure that is dimensionless. Both can be useful in comparing policies.

[5]We also do a calculation on the optimal subsidy and find that the optimal subsidy may be substantially higher than the actual subsidy based on the natural experiment.

[6]This concept could apply to any externality where consumers consider the impact of others' behavior in their decisions.

that for our parameter estimates the shape of the damage function does not appear to have a large effect on our welfare results.

A final set of findings on welfare relates to how subsidies should be targeted. We examined this issue in two ways. First, using our natural experiment, we examined how net benefits per person might change if the most congested route were targeted. We found that targeting users of the most congested route increases average net benefits by 15%. Second, our entire natural field experiment is focused on the issue of targeting. While the field experiment shows that targeting can be effective in shifting demand, the main takeaway is that these shifts may not always increase welfare. A key reason that targeting does not necessarily increase welfare is because the machine learning algorithm used for targeting in the field experiment only takes into account information on crowding. It does not take into account information on unobserved demand characteristics, which are also central for determining the welfare impact of the subsidy. Our experiments suggest that targeting may require a substantial amount of information on economic fundamentals, such as demand, to increase the likelihood that welfare will actually increase with targeting.

Our paper relates to three important research areas. First, our empirical analysis builds on work in sufficient statistics. Our paper is most closely related to Jacobsen et al. (2020). These authors examine how sufficient statistics can be applied to externality-correcting policies; however, they assume markets are perfectly competitive and that consumers take the externality as given, and do not adjust their behavior in response to the level of the externality. We relax those assumptions in Jacobsen et al. (2020), but retain the assumption that consumers are fully informed.[7] Our model also relates to work by Kreindler (2022) on estimating endogenous congestion, but our model differs in terms of its focus on urban mass transit and its theoretical underpinnings.

Second, we build on a large literature on congestion and the welfare effects of peak pricing schemes. It has long been argued that many sectors could benefit from peak-load pricing (Vickrey, 1963; Mohring, 1972; Glaister, 1974; Train, 1977; Glaister and Lewis, 1978; Winston, 1985; Jansson, 1993; Parry and Small, 2009). There have been attempts to estimate the welfare implications of changes in peak or off-peak prices for mass transit (*e.g.*, Parry and Small (2009); Joskow and Wolfram (2012); Knittel and Sandler (2013); Winston (2013);

[7]We do not address issues of estimating welfare with adverse selection, information asymmetries, market structure issues and behavioral agents. For examples of such work, see Bundorf, Levin and Mahoney (2012); Hendren (2013); Weyl and Fabinger (2013); Mahoney and Weyl (2017); Farhi and Gabaix (2020); Ito, Ida and Tanaka (2021). Several studies that focus on externalities and internalites are also related to our work, including Allcott, Mullainathan and Taubinsky (2014); Piketty, Saez and Stantcheva (2014); Allcott and Taubinsky (2015); Allcott, Lockwood and Taubinsky (2019). These studies fit within the general framework developed by Kleven (2021).

Levin, Lewis and Wolak (2017); Kreindler (2022); Tarduno (2022)).[8] Moreover, there appears to be a dearth of well-identified estimates of demand elasticities and cross elasticities for mass transit that have been used to estimate welfare effects. For instance, Parry and Small (2009) estimate the welfare effects of urban transit subsidies using a structural approach, but note that "[l]ittle information is available about shifts of transit riders across time periods" (p. 17). Our empirical analysis helps to address this gap in the literature.[9]

Third, we add to the literature on optimal targeting. Several papers examine how targeting price changes or interventions could increase welfare (Rodrik, 1987; Chassang et al., 2012; Dupas, 2014; Allcott and Taubinsky, 2015; Cohen, Dupas and Schaner, 2015; Alatas et al., 2016; Byrne, Martin and Nah, 2019; Polyakova and Ryan, 2019; Farhi and Gabaix, 2020; Gerarden and Yang, 2021). Our field experiment targeted subsidies using detailed data on congestion and historical consumer behavior, which is typical of many machine learning algorithms in transport (Tizghadam et al., 2019) and other markets (Mullainathan and Spiess, 2017; Athey, 2018; Knittel and Stolper, 2019; Burlig et al., 2020; Christensen et al., 2021; Aiken et al., 2022). However, not having information on price elasticities of demand for travel and consumers' willingness to shift their travel times can limit the effectiveness of targeting subsidies using machine learning. For instance, we show in some time periods that targeted pricing using a machine learning algorithm actually lowered welfare.

The paper is structured as follows: Section 2 provides an overview of the theory that allows us to estimate welfare impacts of the experiment. Section 3 develops demand estimates for the natural experiment and section 4 develops demand estimates for the field experiment. Section 5 implements our welfare analysis. Finally, Section 6 concludes.

2 Overview of the Theory

We develop a theoretical model that allows us to estimate the welfare effects of a subsidy using sufficient statistics. We present the details of our welfare model in appendix A.1. The model identifies key demand parameters of interest that are obtained from our experiments. In

[8] There have been many papers that analyze the impact of peak-time pricing on overall demand and welfare (Borenstein, 2005; Wolak, 2007; Joskow and Wolfram, 2012; Jessoe and Rapson, 2014; Ito, Ida and Tanaka, 2018; Burkhardt, Gillingham and Kopalle, 2019; Fowlie et al., 2021; Williams, 2021). Our paper builds on this work by developing an analytically tractable approach for estimating welfare changes of peak and off-peak pricing in the presence of externalities, where the externalities can affect demand.

[9] Other papers that estimate changes in subway demand with a change in prices include Davis (2021), who estimates the price elasticity of demand for subways in Mexico. Davis' analysis differs from ours in that he focuses on the impact of general price decreases in subway systems, and not on differences in peak and off-peak travel aimed at reducing an externality. Another paper by Yang and Long Lim (2018), estimates the impact of temporary free rides on early morning subway demand in Singapore, and finds a meaningful shift to earlier trains.

addition we consider whether consumers take the behavior of others into account in making their decisions.

The basic model has two types of agents, identical consumers and a firm (e.g., a utility). A representative consumer maximizes utility over peak and off-peak travel, and a numeraire good. We assume that consumer has a well behaved utility function and quasi-linear preferences. In addition to the consumers, a regulated utility is assumed to minimize the cost of meeting demand. Consumer choices generate an externality that affects the utility, which can be thought of as congestion or delay in our application.

We allow prices to deviate from their marginal private cost (which is consistent with pricing in many regulated markets), and we allow for subsidizing particular consumers. The price subsidy is aimed at encouraging off-peak travel in our first experiment (experiment 1), and shifting particular individuals into less crowded trains in our second experiment (experiment 2). The primary aim of the subsidy is to reduce an externality, which can be thought of as crowding on a subway or delay. In our application, we consider both.

The key results from this model are intuitive. The welfare effect of a subsidy depends on how consumers respond to the subsidy, and the extent to which prices deviate from the marginal social cost. If the price is above the marginal social cost in a particular period, and travel increases as a result of the policy intervention (e.g., a subsidy), then it contributes to an increase in welfare. Summing the welfare changes across each period gives the overall change in welfare.

The equation we use to estimate the per capita welfare effect of an off-peak subsidy is given by:

$$\frac{1}{N}(W(s') - W(0)) = (p - MSC_1)\frac{dx_1}{ds}s' + (p - \frac{1}{2}s' - MSC_2)\frac{dx_2}{ds}s' \tag{1}$$

where N is the number of consumers; $W(s')$ is the economic welfare of a subsidy, s'; and $W(0)$ is the welfare associated with no subsidy.[10] $\frac{dx_1}{ds}$ and $\frac{dx_2}{ds}$ are the demand responses to the subsidy in the peak and off-peak periods, respectively. These are assumed to be constant. The price of the good is p, assumed to be constant in both periods. The marginal social cost in periods 1 and 2 are MSC_1 and MSC_2, respectively. Marginal social cost is the sum of the marginal private cost (MPC), which is assumed to be constant across periods, and the marginal external cost, MEC_1 and MEC_2, which can vary across periods.

The derivation of equation (1) makes use of the envelope theorem for the consumer maximization problem. When the off-peak subsidy increases, consumers may alter their travel

[10] We define the change in welfare as the change in producer plus consumer surplus, assuming the subsidy is a lump sum transfer. See appendix A.1 for details.

demand because of an equilibrium change in the subsidy, the price or the congestion level, but the resulting changes in consumption do not a have direct impact on consumer utility. This is consistent with the sufficient statistics approach (Harberger, 1964; Chetty, 2009; Jacobsen et al., 2020).

An important issue for welfare is how consumers respond to the subsidy. We consider the two polar cases of exogenous congestion and endogenous congestion. Exogenous congestion assumes that consumers ignore the impact that the subsidy has on the level of congestion. Endogenous congestion assumes consumers take into account how others will respond to the externality. The two cases can provide insights into the welfare impacts of how interventions scale. The exogenous congestion case may be more appropriate when estimating the welfare impacts of a small-scale subsidy, such as a pilot. The endogenous congestion case may be more appropriate when measuring the the welfare impacts of a relatively large-scale intervention, such as providing a subsidy to all customers.[11]

These two cases will typically result in different values for demand, as measured by $\frac{dx_1}{ds}$ and $\frac{dx_2}{ds}$. Empirical estimates for these cases are presented in section 5. The theoretical relationship between the endogenous case and the exogenous case is discussed in appendix A.16.

3 Natural Experiment

To estimate the welfare impacts of the subsidy, we will need causal estimates of the impact of a subsidy on peak and off-peak travel. This section explains how we derive these estimates for the natural experiment. We provide a brief overview of the Bart network (3.1), a description of the price subsidy and natural experiment (3.2), a discussion of our identification strategy (3.3), and present the key results on peak and off-peak demand (3.4).

3.1 The BART network

The natural experiment was conducted with BART, which is a large public transit system serving the San Francisco Bay area in California.[12] BART is the fifth-busiest heavy rail rapid

[11] This issue applies in many settings, as List (2022) notes.

[12] The heavy rail elevated and subway system connects San Francisco and Oakland with Alameda, Contra Costa, and San Mateo counties across 112 miles of track connecting 46 stations. BART has an average of 423,000 weekday passengers and 124.2 million annual passengers in fiscal year 2017 (BART, 2017),

transit system in the United States. Figure A8 in Appendix B shows the network map.[13] The prices on BART are comparable to those of other U.S. commuter rail systems and are higher than those of most subways, especially for long trips. The minimum price is $1.95 (except for San Mateo County trips) under 6 miles (9.7 km), and the maximum one-way price including all possible surcharges is $15.70, the journey between San Francisco International Airport and Oakland International Airport (2017 prices).[14] Figure A9 in Appendix B plots the distribution of fare by routes. The mean fare for non-airport routes is $4.16, with a median fare of $4.25. The mean and median fare for airport route is $9.60.

3.2 The Perks Treatment

The treatment changed the relative peak to off-peak prices, and Bart called this the *Perks program*. To be eligible for the new pricing program, customers had to have a Clipper ID number.[15] We used Clipper card data for our study. The card records the amount paid, the time the person entered the departing station, and the time they left the arrival station. Linking each user with their Clipper card was necessary to give them the pricing subsidy. This subsidy was based on the frequency, timing, and length of their trips.

The Perks program was advertised over a four week period through direct outreach at stations, media coverage, and employer partnerships. There was some natural time variation in some stations receiving ads before others, and we use this staggered variation later in this section. An example of the ads for direct outreach is shown in Figure A10 in Appendix B. The left hand side of the figure is an example of the basic ad. The right hand side provides a fuller description of the program with eligibility criteria. The program started on August 23, 2016 and ran until the end of February 2017.

The price subsidy changes based on the time the journey starts. For every mile traveled between 630am and 730am or 830am and 930am, the consumer would get a 11.25% reduction in the off-peak price. The peak time at BART was between 730am and 830am. As part of the program to incentivize people to take part, it also gave a modest 2.25% reduction in the peak

[13] BART has five rapid transit lines; most of each line's length is on track shared with other lines. Trains on each line run approximately every 15 minutes on weekdays and 20 minutes during evenings, weekends and holidays; stations on the section of track between Daly City and West Oakland are served by four lines and therefore have 16 trains an hour on each track. BART service begins around 4:00 am on weekdays, 6:00 am on Saturdays, and 8:00 am on Sundays.

[14] The price is based on a formula that takes into account both the length and speed of the trip. A surcharge is added for trips traveling through the Transbay Tube, to Oakland International Airport, to San Francisco International Airport, and/or through San Mateo County, a county that is not a member of the San Francisco Bay Area Rapid Transit District. The farthest possible trip, from Pittsburg/Bay Point to Millbrae, costs less because of the $4 additional charge added to SFO trips and $6 additional charge added to Oakland trips.

[15] Most regular BART customers have a Clipper card, which is a contactless smart card accepted on all major Bay Area public transit agencies, and may be used in lieu of a paper ticket.

price. A relative price difference between peak and off-peak price of 9% was thus created by the Perks program. That is the exogenous change in prices that we will leverage to estimate demand. To receive the subsidy, participants needed to have an active PayPal account that used the same email address as their BART Perks account.

3.3 Identification

Identification of $\frac{dx_1}{ds}$ and $\frac{dx_2}{ds}$ comes directly from this natural experiment. We obtain identification of peak shifting due to changes in off-peak prices using two strategies. First, we compare the self-selected Perks users with the rest of the network using a difference-in-differences framework (we have data on all individual rides on BART before and during the natural experiment). We can determine the share of rides that are in the peak and off-peak hours and use this to obtain a causal estimate of the Perks program on demand. We test for parallel trends in peak and off-peak shares before the Perks program began, and find that the peak and off-peak trends for the Perks and non-Perks users are statistically indistinguishable.

Second, we use the fact that not all BART customers were brought onto Perks at the same time. As mentioned above, some stations advertised the subsidy program before others, so we have quasi-random variation in when people saw the advertisements to the Perks program. The program officially started on August 23th, 2016, but 18.9% of the sample were enrolled in the program after one month from the start of the program. This is our 'late' sample. Comparing 'early' versus 'late' enrollers allows us to estimate the impact of Perks on the number of BART trips, as opposed to just the shares of peak and off-peak travel. Again, we find that the trends for early and late enrollers are quite similar prior to the treatment starting (i.e., they are statistically indistinguishable from each other). We estimate the cross-price elasticity using a difference-in-differences approaches to estimate our welfare model.

We have individual level trip data for the 17,545 BART users who signed up for Perks for the six months before the Perks program began (February 2016) and six months during the Perks program (up to February 2017). This is a total of 2,119,588 individual journeys on the BART network for twelve months. We also have data on the rest of the BART network at 15 minute intervals for the same time period which corresponds to 72,098,160 individual rides for non-Perks customers.

Because prices were not perfectly randomized across BART users, a potential issue with our causal estimates could arise if those who signed up for the program knew that they were more likely to reduce their peak demand during the program. While we cannot rule out this possibility, the near perfect parallel trends in both identification strategies suggest that this is unlikely. To address any concerns, we provide an external check on the results obtained here with our field experiment described in section 4, which randomizes prices across BART consumers. Importantly, we will compare the elasticities from the field experiment to those generated from the quasi-experiments considered here. We show that they are similar.

3.4 Results

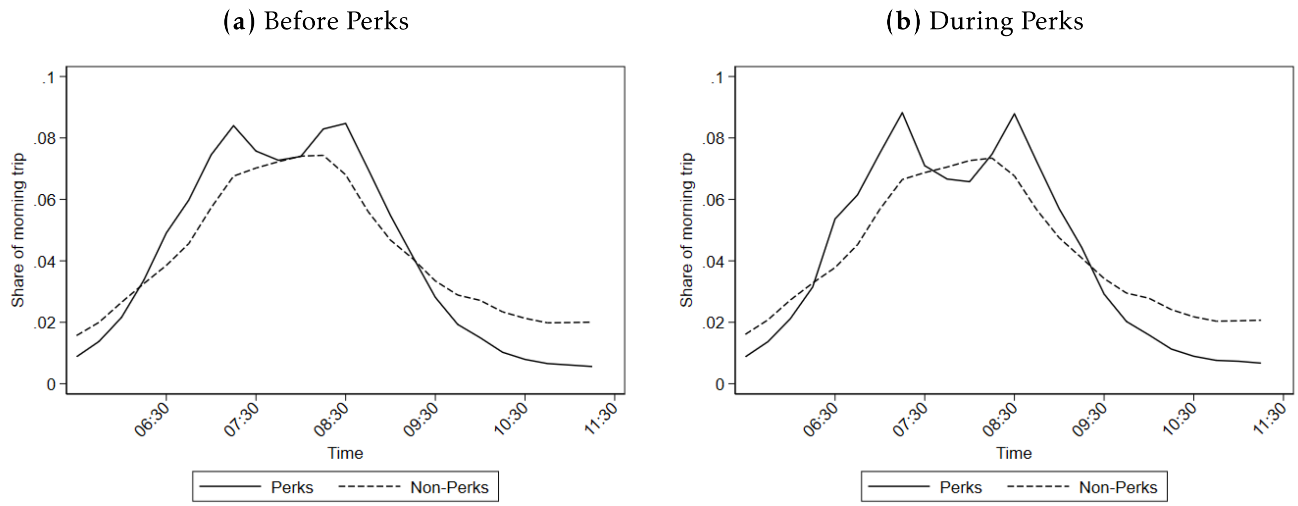
This results section is based on the two identification strategies presented above. Before we present the results, we describe the data of the Perks users and how it compares to the BART non-Perks users. For the rest of the customers using BART, we know what rides (time, day, stations) have been taken in the same time frame as our opt-in sample, but we can not identify individual accounts.

We first compare the distribution of travel demand for our 17,545 BART users in the Perks program with the demand from non-Perks users. Figure 1 below shows the raw distribution of rides taking place in five minute intervals before and during the Perks program. Panel (a) shows demand before the Perks program and panel (b) shows demand during the Perks program. For both panels (a) and (b), the y-axis is the share of trips. The figure reveals that the overall shape of the dashed line for the non-Perks sample is very similar in panels (a) and (b). In contrast, the line for the Perks sample appears to be steeper and higher for the shoulder hours in the morning (i.e., the subsidized hours). This provides some visual evidence that the increase in demand in the shoulder hours from the Perks commuters during the Perks program does not happen with the rest of the BART network customers during the Perks program. This evidence suggests that once customers received the new price changes, they shifted some of their demand to the shoulder off-peak hours from out the peak hour.

3.4.1 Impact of Prices on Demand: Empirical Approach 1

For our first identification approach in comparing the Perks users to the rest of the BART network users, we must ensure that we have parallel trends in the peak and off-peak shoulder hours for the Perks BART users and the non-Perks BART users before the experiment began. This is the necessary condition for our difference-in-differences approach. Figure 2 presents the daily share of peak hour weekday journeys for the Perks and non-Perks users before and during the experiment. This figure shows that there is a very similar trend in the raw peak hour demand data for the Perks users (black line) and the non-Perks users (the gray line) before the Perks program started on August 23, 2018. The trends for the two groups are not statistically significantly different (see the econometric test of pre-trends in peak and off-peak demand in the Perks and non-Perks samples in Figure A12). This allows us to estimate the

Figure 1: Comparing Perks program participants with the rest of the BART network



Note: The figure plots the average daily share of trips taken by the participants in the first experiment in each time interval from 530am-1130am. It also plots on the same figure the average share of daily trips at each time interval taken by the rest of the BART users. The sample for Panel (a) includes April 1, 2016 to August 22, 2017 for all weekdays. The sample for Panel (b) includes August 23, 2016 to February 28, 2017 for all weekdays.

impact of the price change on peak demand using a difference-in-differences framework.

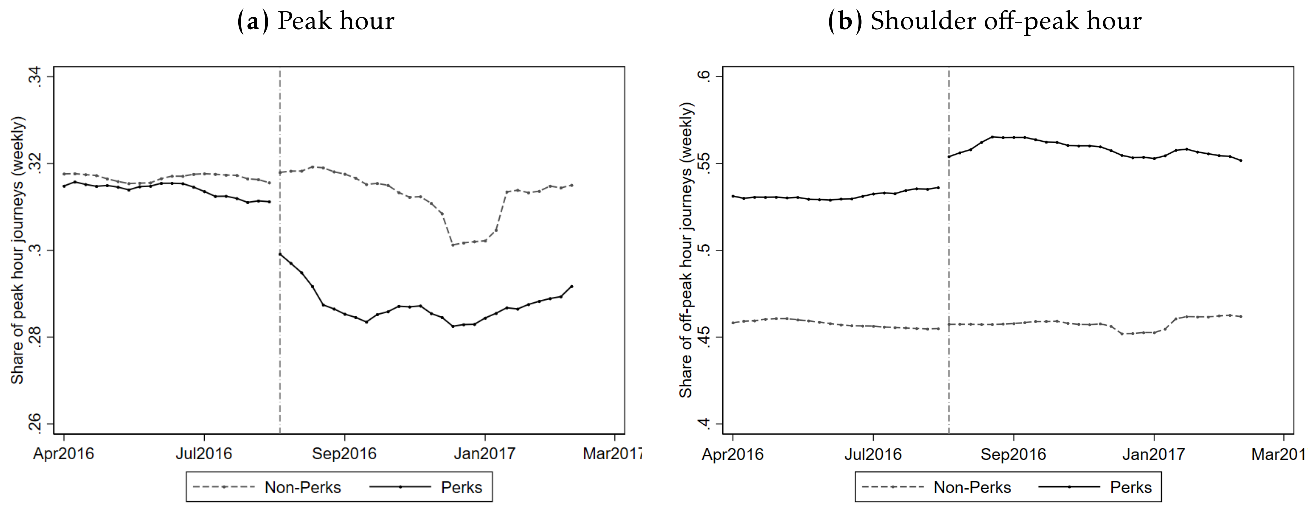
It is also clear from Figure 2 that there is a very similar trend in the shoulder off-peak hour demand data between the Perks users (black line) and other users (the gray line) before the Perks program started. The trends for the two groups are not statistically significantly different before July 2016, as suggested in Figure 2 (see the test of pre-trend in Figure A12). This allows us to estimate the impact of the price change on shoulder demand using a difference-in-differences framework.

Given these two parallel trends in peak and off-peak shoulder hours before the Perks program started for Perks and non-Perks users, we can now estimate the following panel fixed effects model:

$$y_{gt} = \beta_1 * Perks_t + \beta_2 * Treated_g * Perks_t + \beta_3 Treated_g + \gamma_t + \epsilon_{it} \quad (2)$$

where y_{gt} is the share of peak hour (730am - 830am) trips in the morning from 5:30am-10:30am on date t for group g. Treated$_g$ is an indicator for Perks users with the omitted group being non-Perks users. Perks$_t$ is an indicator for the period in which the Perks program is implemented from August 23, 2016 to February 28, 2017. γ_t is a date fixed effect that controls for unobserved changes in travel patterns such as a date-specific service disruption.

Figure 2: Share of peak and off-peak shoulder hour demand for Perks users versus the rest of the BART network users (non-Perks)



Note: Panel (a) shows the five week moving average of the share of peak-hour demand (730am-830am) among all trips from 5.30am-10.30am for Perks and non-Perks users. Panel (b) shows the moving average of the share of shoulder off-peak demand (630am-730am and 830am-930am) among all trips from 5.30am-10.30am for Perks and non-Perks users. The sample includes weekdays from April 2016 to March 2017. The vertical line represents the beginning of the first experiment on August 23, 2016.

Our sample period covers April 1, 2016, to February 28, 2017. It covers five months before the Perks program started and a six month period when the Perks program was running. Our null hypothesis is that the difference-in-difference estimator, β_2, is equal to zero for both the peak period and the off-peak shoulder hour.

Table 1 presents the estimates of equation 2. Column (1) reports estimates on the demand share of peak hour. We find that comparing with other users, Perks users reduce their share of peak hour trips by 2.1 percentage points during the program period. Column (2) reports estimates of the share of off-peak shoulder hours. We find that using non-Perks users as the reference group, Perks users have a 2.6 percentage point increase in the share of off-peak shoulder hour trips during the program period compared to before.

We also estimate an alternative specification that controls for differences between Perks and non-Perks users in terms of the routes they take (defined by the entry and exit station). We compute the weekly share of peak hour (and off-peak shoulder hour) trips for Perks users and for non-Perks users for each individual route in our data This specification allows for some route has sparse number of rides at daily level. We then estimate equation 2 using the weekly share of peak hour (and off-peak shoulder hour) trips for each route and group as outcome, controlling for route-group fixed effects (i.e., the route-specific travel pattern for Perks and

Table 1: Difference-in-difference estimates - Participants and rest of network

Outcome: Share of all trips in morning		
	All dates	
	Peak Hour	Off-peak Shoulder Hour
	(1)	(2)
Treated users × Perks Period (β_2)	-0.021***	0.026***
	(0.0029)	(0.0023)
Treated users (β_3)	-0.0030***	0.0744***
	(0.0008)	(0.0016)
Observations	466	466

Note: Includes daily observations for Perks users and non-Perks users. Sample includes weekdays from April 1, 2016 to February 28, 2017, excluding public holidays. Standard errors in parentheses are clustered by week.

non-Perks users). Table A13 reports the estimates. We find very similar effects as those shown in Table 1.

We next analyze whether the effect of the subsidy applies to the two busiest stations in the BART network, Montgomery Street and Embarcardero. In Table A15 we split the routes into two directions - those that travel from the east of the two stations, thereby crossing the Transbay tube, and those that travel from the west of the two stations. We find that the treatment effect in shifting riders from the peak to the off-peak hour is much larger for the congested routes travelling from the east of the Transbay exiting at Montgomery Street and Embarcardero, compared to those travelling from the west that are less congested.

We also examine the change in demand for peak and off-peak demand when the Perks incentives are removed, which allows us to understand if the incentives created a persistent change in behavior. After six months of the price subsidy program, the program ended. We have BART demand data for those that signed up versus those that did not for six months after the program concluded. We find that while the peak to off-peak shift is lower after the program, there is a persistent and significant treatment effect. We find that the long-term change in peak hour demand without the subsidy was -0.9% (see appendix Table A12). One possible explanation is that people change their demand during the experiment, and once the price changes have been removed, it is costly for users to change their habits or schedules to revert back to their pre-treatment levels. Another explanation is that people may find better ways to commute that increase their welfare, which is supported by evidence from London and Singapore (Larcom, Rauch and Willems, 2017; Yang and Long Lim, 2018).

Despite parallel trends prior to Perks, the customers who selected into Perks may have known prior to Perks that they might want to change their early morning commute time in the future.

This is impossible to test with our data but given how strong the parallel trends are, we find this outcome is unlikely. Another caveat here is that we only observe shares of demand and not absolute demand (i.e., number of trips). This is where the next section and identification strategy helps us.

3.4.2 Impact of Prices on Demand: Empirical Approach 2

For the second empirical approach, we estimated how the Perks price program changed demand in the first month in the program, when 18.9% of users had not already signed up (through the staggered roll-out of the advertisement).We define 'early' enrolled Perks users as those who enrolled in Perks from August 23 to September 2, 2016, where our 'late' control group consist of users who eventually enrolled between Oct 1 to Nov 5, 2016. Therefore in our treatment time period (August 23 to September 31, 2016), all the late enrolled Perks users would have not enrolled in the program yet and thus constitute an effective control group for the travel behavior in the absence of the Perks program.[16]

A benefit of using early versus late Perks users is that we can analyze absolute changes in demand at the individual level as a result of the price change. Moreover, both early and late customers all enrolled into the price change, so unobserved characteristics are even less likely to drive the change in peak demand. In addition, the approach of comparing an early treatment group to a late control group is suggested by Callaway and Sant'Anna (2021) as a way to overcome some of the issues in two-way fixed effects models. Figure A6 shows the cumulative distribution for BART enrollment as a function of time.

We can analyze whether the pre-Perks trend before the Perks program is the same for the early and late enrollers. Figure 3 suggests that the trend on peak and shoulder hours before the Perks program is the same. In Figure A13 we test for the difference of the pre-trend by each of the week, we find that for peak hour trips, the pre-trend for the two groups are not statistically significant. For off-peak shoulder hour trips, we find no systematic difference in the trend for the two groups as well.

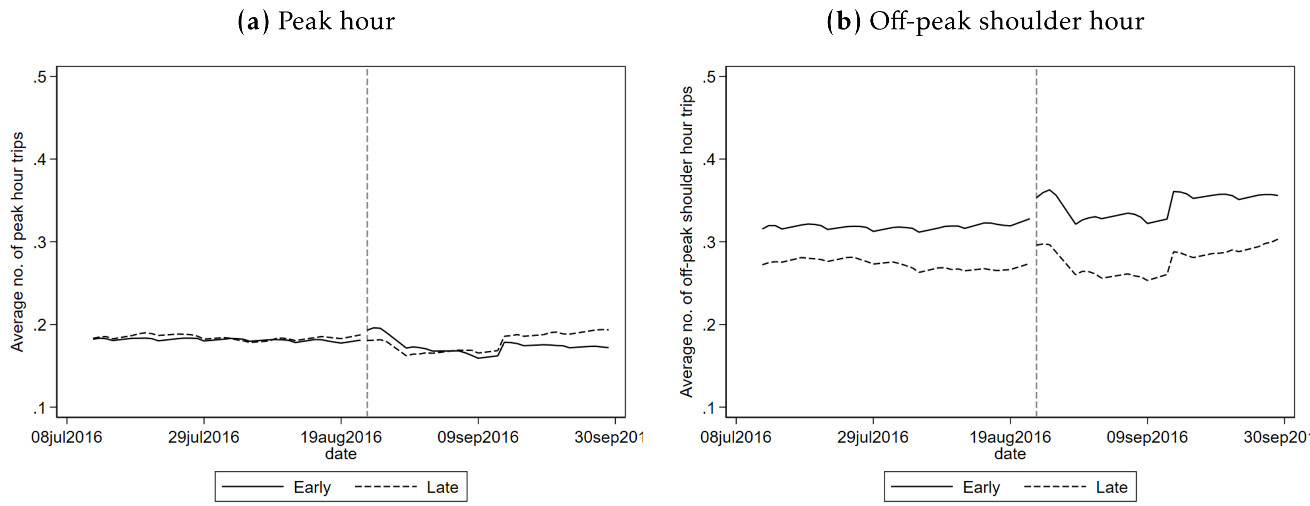
We can now estimate the following equation:

$$y_{it} = \beta_1 * E_i * Perks_t + \sigma_i + \mu_t + \epsilon_{it} \tag{3}$$

where y_{it} measures the number of trips during Peak hour on date t for user i. E_i is an indicator for individual i being enrolled early from August 23 to September 2, 2016. $Perks_t$ is an

[16] Our analysis is robust to the time period selected for early versus late. We present in Appendix B Table A14 results that define late enrolled users as those enrolled between Oct 15 to Nov 5.

Figure 3: Early and late enrolled



Note: The sample includes trips from July 12, 2016 through September 30, 2016. The figure plots the two week moving average for off-peak shoulder hour trips (630am-730am and 830-930am) and peak hour trips (730am-830am) of early enrolled and late enrolled users. The vertical line represents the date August 23, 2016, when the Perks program started.

indicator for the perks program period. σ_i is an individual fixed effect and μ_t is a date fixed effect. Our null hypothesis is that the difference-in-difference estimator, β_1, is equal to zero in both the peak hour and the off-peak shoulder hour.

Table 2 presents the estimates for equation 3. Columns (1)-(2) report the estimates on the number of trips per day in the time window of peak hour and off-peak hours respectively. We find that compared with the late enrolled users, during the first month of the program the daily number of trips during the peak hour by early enrollers decreases by 0.007; for the off-peak shoulder hours the daily number of trips by early enrollers increases by 0.015. Column (3) reports the estimates on the total number of trips in the morning.[17] We find that the Perks program does not increase the total number of trips taken by early enrollers.

To convert these numbers to percentages, we need to know the baseline number of trips. The baseline number of trips for the late enrollers during September is 0.178 and 0.273 during the peak and shoulder hours respectively. The DiD estimate of -0.00705 for the peak hour represents a 4% decrease in peak demand. The DiD estimate of 0.0151 for the off-peak shoulder hours represents a 5.5% increase in off-peak shoulder hour demand.

[17]The estimates correspond closely to the estimates in share in Table 1. The share of peak (off-peak) trips implied among all trips from 5.30am-10.30am is 0.303 (0.535), with reduction of trips of -0.007 (increase of 0.0151), and the share of peak (off-peak) trip during the program would be 0.284 (0.566), which is an implied change of shares of -0.019 (0.0304).

Table 2: Average treatment effects by time period: Early vs. late comparison

	OLS		
	Outcome: Number of trips daily		
	Peak	Off-peak shoulder (Either)	Total (5.30-10.30am)
	(1)	(2)	(3)
Early enrolled users × Perks Period (β_1)	-0.0071**	0.0151***	0.0040
	(0.0031)	(0.0038)	(0.0044)
Observations	600148	600148	600148
Date FE	x	x	x
User FE	x	x	x
Sample mean (before Perks)	0.18	0.31	0.57

Note: The sample includes daily trips from June 1 to October 1, 2016, excluding August 23 to September 2. The Perks period is defined as dates after September 2. Early enrollers are defined as those who enrolled between August 23 and September 2. The control group includes those who enrolled between October 1 and November 5, 2016. Standard errors are clustered at the user-week level. The sample means are for late enrolled users in periods before the Perks period.

To check whether the treatment effect is robust to different specifications, we relax the specific definition of early versus late enrollment. Instead, we make use of the fact that the activation time into treatment for each user differs. We estimate the change in travel behavior of each user before and after their Perks program activation using the following equation:

$$y_{it} = \eta_1 PerksActive_{it} + \mu_t + \sigma_i + \epsilon_{it} \tag{4}$$

where $PerksActive_{it}$ is an indicator that turns on when user i activated their Perks status on date t. σ_i are user fixed effects and μ_t are time fixed effects. As above, our null hypothesis is that the difference-in-difference estimator, η_1, is equal to zero in both the peak hour and the off-peak shoulder hours.

Table 3 presents the estimates for equation 4 and shows that the effects of the price subsidies on peak and off-peak demand are very similar to our staggered specification. We find that the after activating their Perks user status, peak hour journeys decreased by 0.007 per day for each user and bonus hour journeys increased by 0.024. The latter result is significantly larger than that from the early versus late enrolment specification. We also estimate the treatment effect following De Chaisemartin and d'Haultfoeuille (2020) and Callaway and Sant'Anna (2021) and report the estimates in Appendix Table A18. Using these approaches does not alter our estimates, providing more confidence in our identification strategy.

Table 3: Average treatment effects by time period: staggered activation

	OLS		
	Outcome: Number of daily trips		
	Peak	Off-peak shoulder (Either)	Total (5.30-10.30am)
	(1)	(2)	(3)
Active status level	-0.00713***	0.0244***	0.0130***
	(0.00221)	(0.00389)	(0.00454)
Observations	1426285	1426285	1426285
Date FE	x	x	x
User FE	x	x	x
Sample mean (before Perks)	0.17	0.30	0.56

Note: Sample includes dates from July 1, 2016 through November 6, 2016. Standard errors are clustered by users and dates.

3.5 Own and cross-price elasticity of demand for peak and off-peak travel

Given our estimates above, we can now estimate the own and cross-price elasticity of demand for peak and off-peak travel. We estimate $\hat{\eta}$ from equation 4 from estimating the demand effects in our two difference-in-difference specifications. We estimate the causal impact of changing relative prices on the demand for rides taken per day in peak or off-peak periods (we denote η_j as the estimates for the equation where the outcome is travel in period j).

We define the relative price of off-peak, p_2, as the ratio of price per mile of off-peak to that of peak. The estimated marginal price of a mile demanded on BART is \$0.112 in the control condition, where there is no difference between peak or off-peak prices. With no price subsidy, this relative price of off-peak is $p_2 = \frac{0.112}{0.112} = 1$, and it applies to the control group. For the treatment group, the price of one mile on BART is reduced by 9.2%. We calculate the average off-peak shoulder hour trip taken per day by rider before the program as $\bar{x_2}$. Our estimate on own-price elasticity of off-peak hour equals to $\epsilon_{22} = \frac{p_2}{\Delta p_2}\frac{\eta_2}{\bar{x_2}}$ Our estimate on cross-price elasticity of peak hour demand to off-peak hour price equals to $\epsilon_{22} = \frac{p_2}{\Delta p_2}\frac{\eta_1}{\bar{x_1}}$. Our estimates for own-price elasticity of off-peak demand is -0.86, and the cross-price elasticity of peak demand with respect to off-peak price is 0.44.[18] We find that this own-price elasticity does not meaningfully vary by whether we examine consumers who shift to the early off-peak hour as opposed to the later off-peak hour (see appendix Table A16).

[18]The program offered 5 points per mile of off-peak trip travelled. Each point has monetary value of \$2.52/1000, therefore the subsidy provides a monetary value of \$5*2.52/1000 = 0.0126 per mile in off-peak. The subsidy also gives 1 mile per peak trip travelled. This implies a new relative price of off-peak as $\frac{0.112-0.0126}{0.112-0.00252} = 0.9079$. Therefore, the relative price change is $\frac{\Delta p_2}{p_2} = -0.0921$. From Table 3, the estimate for η_2 is 0.0244. The average off-peak trip per day is $\bar{x}_2$ is 0.307, therefore, $\epsilon_{22} = -0.86$ (s.e. = 0.137). As $\bar{x_1} = 0.176$, and η_1 from Table 3 is -0.00713, $\epsilon_{12} = 0.44$ (s.e. = 0.136).

We find that the cross-price elasticities of demand change in ways that one would think based on economic theory. For example, the closer the consumer typically travels to the off-peak hour before the subsidy, the greater their own-price response (see appendix Figure A14).

We also find that the more congested the network, the greater the cross-price elasticity (see Table A15 in the appendix). For example, we examine the shift for any route exiting at Montgomery Street and Embarcadero (the two most congested stations of the BART network). Demand from the east (i.e., eastbound Transbay tube trips that are in the busy corridor) towards these two stations shift more (change in absolute share) than that from the west (i.e., westbound transbay tube trips). While all of these heterogeneous effects go in the direction that we expect, they are still correlations and not the causal estimates of changing these parameters on the cross-price elasticities of demand. Nevertheless, we can estimate welfare changes with these different samples in the BART network.

4 Field Experiment

In this section, we estimate the change in peak and off-peak demand from the field experiment. We use these demand estimates to develop a causal estimate of the impact of the subsidy on welfare in the next section. Here, we provide a brief overview of the field experiment and a discussion of our identification strategy, and then present the key results on demand. We then compare the estimates with those obtained in experiment 1.

4.1 Identification

The identification of price elasticities in this field experiment comes from different variation than in the natural experiment. First, we control the assignment mechanism of price changes in this field experiment, whereas we did not control assignment in the natural experiment. The clear advantage of this experiment was the ability to randomize some BART customers into the control group where prices did not change and some BART customers into the treatment group where prices were changed.

Second, in contrast to the natural experiment, which focused on generally shifting riders from peak (730am-830am) to off-peak (630am-730am and 830am-930am) travel, this field experiment focuses on shifting riders to less crowded trains in time periods adjacent to their normal commute. For example, a rider may receive a subsidy to shift their travel from their usual time period to an adjacent twenty minute interval, when a less crowded train is available at their departing station.

Third, in comparison with the natural experiment, this field experiment differed in terms of how price changes were determined. Prices were customized based on the individual's previous BART demand and location. Prices were also updated at least monthly based on changes in congestion levels and other factors. In contrast, the natural experiment offered the same incentives to all participants for the duration of the pilot.[19]

Overall, the general size of the incentives, the way they are paid to consumers, and the type and selection of consumers are very similar across both experiments. The price changes started in mid-December 2018 and ended in March 2019. We could not examine any persistence in travel behavior in this experiment as those in the control group received offers in April 2019. Customers registered for the price subsidies via BART's website and the BART Official Mobile Application.

In terms of defining our field experiment, we class it as a hybrid between a natural and a framed field experiment, as defined by Harrison and List (2004). While we have selection of people into the experimental sample who wanted some incentives, people did not know that they were part of an experiment and did not know that they would get price subsidies for changing their commute time.[20] Therefore, we believe that there is little selection on peak and off-peak price elasticities due to the opaque advertising in the field experiment about what incentives would be given.

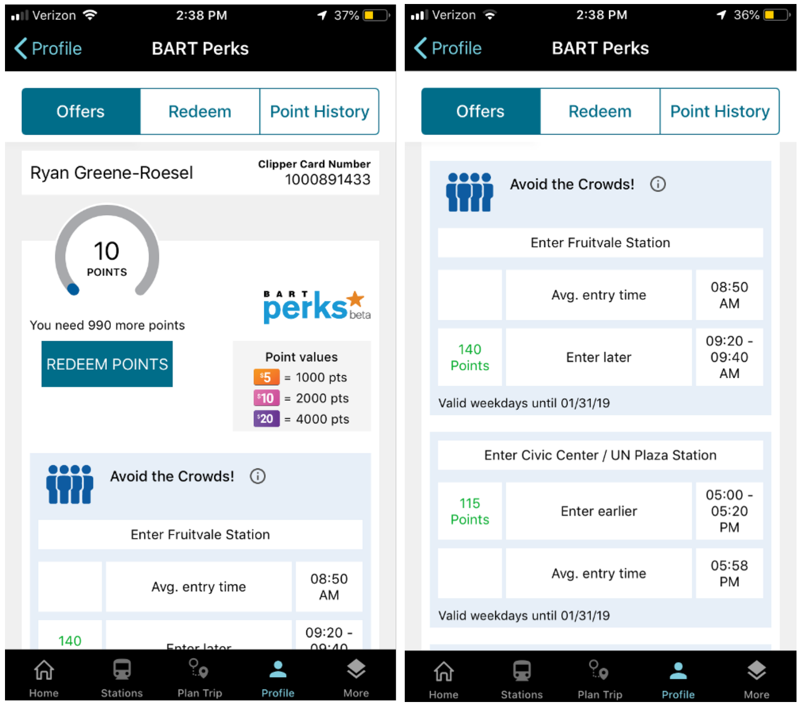
4.2 Recruitment

Participants were recruited from two sources. First, twelve people employed by BART distributed flyers advertising that people could sign up for the program at the Embarcadero, Montgomery and Civic Center stations in downtown San Francisco from 8:00 – 9:30 AM on December 13th, 2018, These three stations were chosen because they are the most congested stations in downtown San Francisco. Second, a subset of those who participated in Perks natural experiment opted in to receiving notifications from BART about future incentive programs. We analyzed the travel histories of these individuals, and identified customers who made at least 50 percent of their BART trips from downtown Embarcadero, Montgomery, or Civic Center stations. An email invitation to join the Perks Phase II program was sent to a random subset of qualifying individuals. Emails were sent in batches between November 20 and December 12, 2018. The overall recruitment led to 1900 customers, of which 63% were recruited through flyers and 36% through email. See appendix Table A19 for the summary statistics for our sample.

[19]Price subsidies were paid out through gift cards as opposed to PayPal (as in the natural experiment).

[20]The program was advertised as giving monetary payments for completing surveys and traveling on the BART on evening and weekends. See the recruitment flyer in Figure A15.

Figure 4: View of the incentives offer on the app



4.3 Price changes

Experimental subjects in the treatment group were offered a price reduction of around 25% per ride in an adjacent time window to their current commute time. The price subsidy had two components: a random term and a deterministic term. The latter was based on how far the time interval was from their usual departure time. The price changes were targeted to each customer's frequently-visited station at a specific 20-minute time window during the morning and/or evening commute period. The time window where the subsidy was offered was up to 40 minutes before or after their typical entry time, and was optimized to reduce overall crowding on the BART network. Users could receive up to four commute-related point offers (shift early AM, shift late AM, shift early PM, shift late PM).

The price subsidies were determined by a machine learning algorithm that drew upon a participant's travel history and a predictive model for crowding provided by the transit company Metropia. Figure 6 provides an example of an offer to shift commuting time. The price subsidies were offered in terms of points that were to be redeemed by the consumer.

To identify the correct time windows to reduce overall BART congestion, the Metropia algorithm identified a user's typical departure time and then calculated the predicted crowding reduction benefit of shifting this user to one of the adjacent time periods. If no benefit would occur from shifting the user, then no offer would be shown. Similarly, no offer was shown if achieving the crowding reduction would require the user to make shifts beyond 40 minutes from their average departure time.[21] The consumer was emailed the price change at the beginning of the month and the price change was constant for one month, so we are estimating one-month elasticities in this field experiment (in comparison with six month elasticities from the natural experiment). During the course of the field experiment, about $23,000 in peak shifting price subsidies were given to BART consumers.

4.4 Results

Our analysis of the field experiment consists of two parts. First, we provide the overall treatment effects of the incentives on shifting ride demand later or earlier. Second, we provide a more precise estimate of the welfare effects of shifting travel into or out of specific 20 minute time periods.

To estimate the average treatment effect of the price subsidies for shifting demand earlier or later than the usual travel time, we employ a linear regression of whether a rider travels at the subsidized time period based on the subsidy they receive as the dependent variable.

For the shift early subsidies, we estimate the following equation:

$$y_{it}^{SE} = \alpha + \beta T_{it}^{SE} + \epsilon_{it} \tag{5}$$

Here, y_{it}^{SE} is an indicator variable for rider i who receives the shift early subsidy and travels at the subsidized time on date t. T_{it}^{SE} is an indicator variable for rider i in the treated group who receives the subsidy (while it is zero for control group riders who receive a hypothetical subsidy). The parameter β captures the overall effect of the shift early subsidies on travel at the subsidized time.

Similarly, we define another indicator variable u_{it}^{SE} to capture whether a rider received a shift early offer but still travels at their usual travel time on date t. We estimate the following equation:

$$u_{it}^{SE} = \gamma + \delta T_{it}^{SE} + \mu_{it} \tag{6}$$

Here, u_{it}^{SE} represents the indicator variable for rider i who received a shift early offer but still travels at their usual time on date t. T_{it}^{SE} is the treatment indicator, indicating whether rider i received a shift early offer on date t. The parameter δ measures the average effect of the subsidy in reducing travel at the usual time for riders.

[21] This interval was based on feedback from user focus groups, which found that individuals did not wish to make large shifts in their commute

We follow a similar approach to estimate the effects of the shift late subsidies, using indicator variables y_{it}^{SL} and u_{it}^{SL} to capture riders traveling at the subsidized time and their usual time, respectively.

In Figure 5, we present the likelihood of riders traveling at their usual time or the subsidized time, comparing the treatment and control groups. The results show a clear effect of the subsidies in shifting riders from their usual travel time to the subsidized time. Specifically, for the shift early subsidies, the estimated equation (5) suggests that the subsidies increased the probability of riders traveling at the subsidized time (compared to the control group) by 0.034 (s.e. 0.009), or 3.4%. Equation (6) suggests that the subsidies reduced the probability of riders traveling at their usual time by 0.035 (s.e. 0.016), or 3.5%.

For the subsidies that shift riders to a later time, the analysis indicates an increase in the probability of riders traveling at the subsidized time by 0.022 (s.e. 0.008), or 2.2%, and a reduction in the probability of riders traveling at their usual time by 0.046 (s.e. 0.017), or 4.6%. These estimates demonstrate that the subsidies successfully shifted transit demand from congested time periods to non-congested time periods, resulting in meaningful and significant changes in travel behavior.

Our second part of analysis of the field experiment leverages the detailed part of the experiment, where we have varying off-peak and peak time windows for each consumer, so a flexible specification is required to account for this. We will also turn these estimates into welfare calculations. We estimate the following equation:

$$y_{it}^{k} = \beta_{k1} S_k + \beta_{k2} S_{-k} + \delta_{kjt} + \delta_{-kjt} + \mu_{ia} A_t + \mu_{id} + \mu_{iq} + \gamma_t + \epsilon_{it} \tag{7}$$

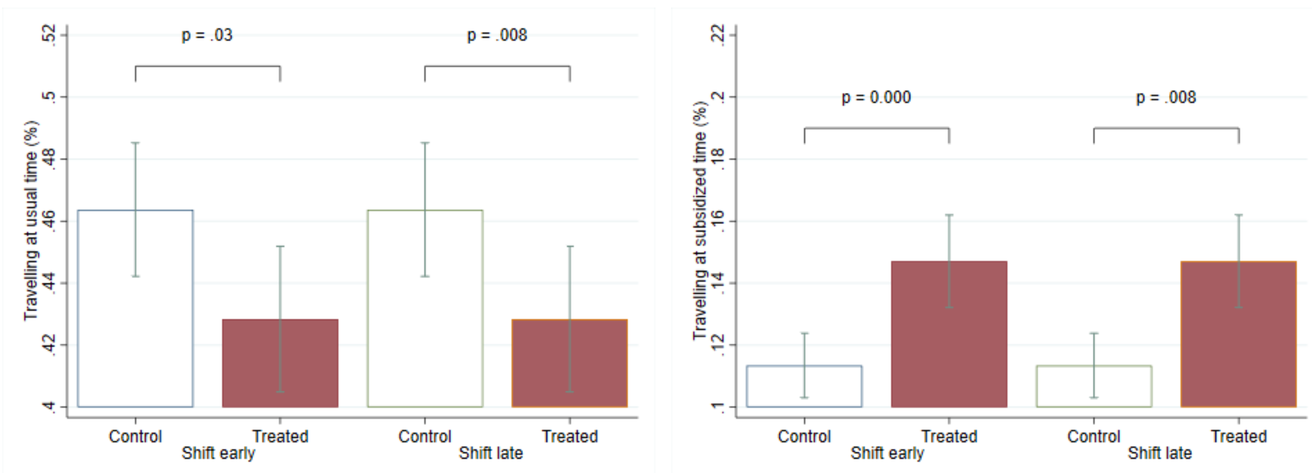
where y_{it}^{k} is the number of trips that start in time period k by individual i on date t. Our sample includes both the treatment and control groups, and both time periods (before and during the experiment). S_k is the monetary value of the points subsidy offered to individual i for time period k, regardless of the "usual travel time" of individual i (e.g., if k is 8am, S_k is any subsidy for 8am, whether it is for someone who usually travels at 7.20am, 7.40am, 8.20am or 8.40am). $S_k = 0$ for the treated group in the pre-treatment period, for the control group, and for those who are not offered a subsidy for time period k. β_{k1} is therefore the own-price demand effect for time period k.

S_{-k} is the value of the subsidy received by individual i for any time period outside of time period k, targeted for someone who usually travels at time k. It is the value of the subsidy offered to individual i who usually travels at time period k that may shift them away from time period k (e.g., if k is 8am, S_{-k} are subsidies for 8.20am/8.40am/7.40am/7.20am that shift people away from 8am). For individuals who do not normally travel in time k, $S_{-k} = 0$

Figure 5: Field Experiment: Change in demand for subsidized time and usual travel time

(a) Travel at usual more congested time

(b) Travel at subsidized less congested time



Note: The figure presents the share of days riders travelling at their usual time (Panel (a)), and the subsidized time (Panel (b)), by the treatment and control riders, and by the period they receive shift early or shift late subsidies. Each bar represents the share of days the outcome (travelling at the usual time or the subsidized time) takes the value of one for the respective group. The red bar(s) represent the average outcome for treated riders and the white bar(s) represent that outcome for the control riders for the respective groups. The p-values above the pairs of bars indicates the statistical significance for the differences between the treated and control riders in the outcome variable, estimated by the regression equations (the treatment coefficients are 0.0336 (0.00928), 0.0217 (0.0082), -0.0353 (0.0163) and -0.0458 (0.0171) respectively for the shift early usual time, shift late usual time, shift early subsidized time and shift late subsidized time). Standard errors are clustered at the rider level and are in parentheses. The sample includes dates when riders receive a shift early (or late) offer or a hypothetical shift early (or late) offer for the control group riders. The experimental period begins December 2018 and goes through March 2019.

for any subsidy they were offered. k.[22] β_{k2} is therefore the generic cross-price effect on the demand for time period k when a price subsidy was offered in time periods other than k.

δ_{kjt} is a time-varying fixed effect for the individuals who normally travel in time j and have received a subsidy for time period k (i.e., those who were actually chosen by the Metropia machine learning algorithm to travel on a less congested train within a 40 minute window of their usual travel time). Each type of rider, jk, includes riders from the treated group and the control group. For the treated group it is defined by the price subsidy offered (those who normally travel in time j who received a subsidy for time period k), and for the control group, it is defined using the hypothetical offer that would have been given to the control group if they were in the treatment group (those who normally travel in time j who received a hypothetical subsidy for time period k).[23] The fixed effects therefore control for any time-varying factors that affect the travel pattern of similar types of riders in the treatment and control group. For the control group, $S_k = 0$. Hence, the identification of S_k comes from a comparison of riders of similar "types" as defined by Metropia offers in both the treatment and control groups.[24]

δ_{-kjt} is the time-varying fixed effect for the individuals who normally travel in time k and were offered a subsidy for time period j. Each type of rider, jk, includes both riders from the treated and control groups. For the control group, $S_{-k} = 0$. Hence, our identification for S_{-k} comes from a comparison of riders of similar type in the treatment and control groups that were identified as having similar offers.

μ_{id} is a fixed effect for individual-day of the week. It controls for individual travel pattern on each day of the week. μ_{iq} is an individual*year quarter fixed effect, which allows for an individual's travel pattern to change in each quarter. μ_{ia} is an individual fixed effect that takes a value of one for the weeks beginning November 12 and 19, 2018, in which the air quality in San Francisco was a historically poor level due to wildfires. It controls for any idiosyncratic response of each individual during those two weeks. γ_t is the date fixed effect. Our null hypotheses are that the price subsidy in the targeted time period has a zero impact on demand and that the cross-price effect on demand is zero.

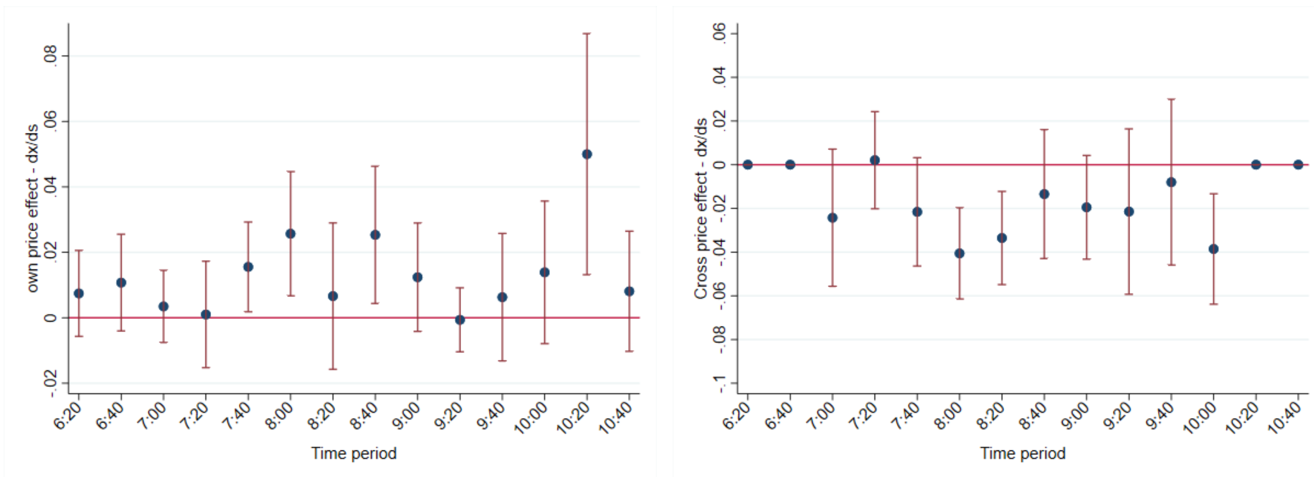
[22]Specifically, let S_m^k be the subsidy that a targeted traveler who usually travels in time period k and subsidize travel in time period m. As in our setup the subsidy shift traveller earlier and later by 20 or 40 minutes interval, if we index the time interval according to the order of time of the day, $S_m^k = 0$ if $|k-m| > 2$. S_{-k} is defined as $S_{-k} = S_{k-2}^k + S_{k-1}^k + S_{k+1}^k + +S_{k+2}^k$. As each traveller has a specific usual travel time k, and receive subsidy for only one time period other than k at each period t, it is equivalent to $S_{-k} = \max\{S_{k-2}, S_{k-1}, S_{k+1}, S_{k+2}\}$.

[23]The algorithm created hypothetical offers for all consumers, but we only gave the offers to the treated group and the control group was unaware of such offers.

[24]Equivalently, these are the fixed effect coefficients for a time invariant indicator D_{kj} interacted with time dummies, where the indicator D_{kj} is 1 if the rider i in the treatment group has ever received subsidy of type S_{kj} in the experiment period, or if the rider i from the control group has ever received a hypothetical subsidy of type S_{kj}.

Figure 6: Field Experiment: Demand Parameter Estimates

(**a**) Change in subsidized time (**b**) Change in usual time



Note: The figure plots the estimates of equation 7. The sample include participants in the field experiment. Panel (a) plots the estimate for the effect of a small change in the subsidy at time t on the demand at time t. Panel (b) plots the estimate for the effect of a small change in subsidy for time t, where the usual travel time is time j, on the demand at time j. Standard errors are robust to autocorrelation of 14 days. The average journey price in the control group is $3.99.

Figure 6 plots the demand estimates, β_{k1}, for each of the time periods in equation 7. Panel (a) plots the demand change in the time period that subsidies were offered. We find, in general, that the estimates for βs are positive, meaningful, and statistically significant, suggesting riders are responsive to price at relatively narrow time intervals (as compared to the wider window in experiment 1). The time periods with the greatest demand shifts toward that period are 740am-8am, 8am-820am, 840am-9am, and 1020am-1040am. The average own-price elasticity of demand from peak to off-peak is -.939 . However, there are some time periods where people are less elastic in their demand, such as before 740am and after 9am (excluding the 1020am-1040am time period).

Panel (b) plots β_{k2} of equation 7, which is the change in the usual travel time for the consumer. In general, we find negative and statistically significant coefficients, suggesting that riders were reducing travel in their usual time and shifted into the subsidized time period (panel (a)). The average cross-price elasticity of demand from peak to off-peak is .543 .[25] It seems

[25]The cross-price elasticity for each type of treatment subsidy (a subsidy for those who normally travel in period j' receiving a subsidy for time period k') is calculated as $\epsilon_{j'k'} = \frac{dx_{j'}}{dS'_{-j}} * \frac{p}{x'_j}$ where $x_{j'}$ is the average travel (in number of trips) for those who normally travel in time period j' who received subsidy in time period k', $\frac{dx_{j'}}{dS_{-j'}}$ is the estimate of the coefficient for S_{-k} in equation (7) for the time period $k = j'$, and $p = 3.99$. The average elasticity is arithmetic average of the elasticities $\epsilon_{j'k'}$all the treatment subsidies (include both shifting early and late).

that consumers are more elastic between 7am and 10am, but less elastic before 7am and after 10am.

From both panels in Figure 6, it is clear that elasticities vary across time periods. This variation in elasticities results from a combination of individual preferences and BART system effects. For example, the own-price and cross-price elasticity at 8-820am are -1.45 and 0.30 respectively.

For those individuals who did not switch at any time period, we sent them a survey asking them why they did not switch earlier or later.[26] We found that for those who would not switch earlier, the most chosen reasons were personal preference (55%) and employer would not allow it (29%), and that for those who would not switch later, the most chosen reasons were also personal preference (41%) and employer would not allow it (40%). This suggests possible unobserved reasons for not switching demand to an earlier or later time period.

The previous analysis allows us to compare the elasticities from the field experiment to those from the natural experiment. The own-price elasticity of demand and cross-price elasticity of demand from the natural experiment were -0.86 (se = .14) and 0.44 (se = .14) respectively. Given the respective estimates from the field experiment are -0.94 and 0.54, we cannot reject that the estimates from both experiments are equal for both the own-price elasticity and the cross price elasticity.

5 Welfare analysis

In this section we use the theory presented in section 2 and estimates of demand from the natural and field experiments to derive the change in welfare. We also present an alternative welfare calculation using the marginal value of public funds (MVPF).

We present an overview of the key data used in the calculation in section 5.1. We then present welfare estimates for the natural experiment in section 5.2. We extend this analysis to allow for differences in demand across groups (in this case, customers that received or did not receive a subsidy) in section 5.2.2. Finally we estimate the MVPFs for the natural experiment in section 5.3 and the field experiment in section 5.4.

The welfare analysis yields several novel findings. First, providing the off-peak subsidy for BART users improves welfare both when we use our welfare model and when we compute the MVPF (Finkelstein and Hendren, 2020). We find that the average net benefits per dollar of direct subsidy is \$0.36 in the base case. We also show that the MVPF in this case is 1.6. This means that for each additional dollar of net cost to the government, the marginal benefit

[26] Survey implementation and results can be found at BART Perks Phase II Evaluation Report (2019).

to all parties is \$1.60.[27] The increase in welfare from the subsidy results from an increase in overall ridership and a shift in ridership from the peak to the off-peak period, which reduces congestion.

Second, allowing for riders to take account of how other riders respond to congestion (*i.e.*, the case of endogenous congestion) also leads to an increase in welfare, but less of an increase than for the case of exogenous congestion. The reason for a higher level of welfare in the case of exogenous congestion is that fewer people ride in the peak period (which yields a smaller welfare loss), and more people ride in the off-peak period (which yields a larger welfare gain).

Third, based on the natural experiment, targeting particularly congested routes can increase average per capita welfare. We estimate that per capita welfare gains increase by 15% when targeting subsidies to a highly congested route—the Transbay route.

Fourth, the optimal subsidy may be substantially higher than the subsidy offered in the natural experiment. We estimate that the optimal subsidy could be about eight times higher than the subsidy that was offered, and that welfare could be about four times higher. As we note below, the particular numerical results are likely to be sensitive to our assumption of linearity of demand.

Finally, using the natural field experiment, we show that the targeting of subsidies to time periods that are relatively narrow can increase welfare, but does not always do so. Our experiments suggest that such targeting may require a substantial amount of information on demand to increase the likelihood that welfare will actually improve with a change in the level of the subsidy.

5.1 Natural Experiment: Data

We analyze the welfare effect of experiment 1 using equation (1) in section . We consider both the case of exogenous and endogenous congestion. The key parameters used in equation (1) in the "base case" are presented in Table 4.

The expressions for $\frac{dx_1}{ds}$ and $\frac{dx_2}{ds}$ (the impact of a change in the subsidy on peak and off-peak demand) are obtained from Section 3 for the case of exogenous congestion.[28] These quantities are measured in terms of the change in the average number of trips per weekday with a change in the subsidy. Price is obtained by averaging the fare for all BART trips in our sample, measured in dollars per trip. This includes all trips in the morning from 5am to noon, excluding those that begin or end in the airport because they have a different fare schedule.

[27] We compute all MVPFs based on the assumption that that the initial subsidy is zero.

[28] For endogenous congestion estimates, see the discussion in appendix A.6 and A.16.

Table 4: Key parameters used in our base case welfare analysis

Parameter	Exogenous congestion	Endogenous congestion	Percent Change
$\frac{dx_1}{ds}$	-0.0194	-0.0166	-14.68
$\frac{dx_2}{ds}$	0.0664	0.0625	-5.91
p	\$3.99	\$3.99	-
s'	\$0.25	\$0.25	-
MPC	\$1.89	\$1.89	-
MEC_2	0	0	-
MSC_1	\$3.22	\$3.22	-
MSC_2	\$1.89	\$1.89	-

Notes: Demand decreases in the peak period and increases in the off-peak period with the introduction of the subsidy. The magnitude of the decrease is smaller with endogenous congestion in the peak period, and the magnitude of the increase is smaller with endogenous congestion in the off-peak period. See discussion in the text for sources of the parameter estimates. Demand is measured as the average number of trips taken per weekday for the average rider.

The subsidy, s', is taken from the actual experiment and equals \$0.25 per trip on average. MPC refers to marginal private cost of a BART trip. MSC_1 and MSC_2 refer to the marginal social cost per trip in the peak and off-peak period, respectively.

The marginal private cost in each period is assumed to be the same in both periods. We estimate the MPC from BART's financial report from 2006-2016. We follow Winston and Maheshri (2007), and regress yearly total operating cost (adjusted for inflation) on yearly annual ridership to obtain the marginal private cost. In addition, we assume that there is no externality in the off-peak period, so the marginal external cost in the period, MEC_2, equals zero. This implies that the marginal social cost in the off-peak period equals the marginal private cost.

To obtain the marginal social cost in the peak period, we need an estimate of marginal external cost (MEC_1) to add to the MPC. The estimate of MEC_1 is based on crowding, delay and the value of time. Our analysis assumes that external costs are proportional to the value of time for both crowding and delay (see Haywood and Koning (2015)). We also assume that MEC_1 is constant for given values of the value of time, congestion and delay. We assume the value of time is half the hourly wage (Small, 2012), also do a sensitivity analysis with 75% of the wage rate (Goldszmidt et al., 2020). Congestion and delay are derived from BART data (see appendix A.4 for details of this calculation). Using this approach, we find that MEC_1 is \$1.33 per peak hour trip.

Table 4 shows that the assumption about the response to congestion matters for the demand response to the subsidy in both the peak and off-peak periods. Even though peak travel declines with the introduction of the subsidy in both cases, the decline is about 15% less with endogenous congestion than with exogenous congestion. More people travel during the peak period when they take into account how others will respond to congestion. Similarly, fewer people travel during the off-peak period with endogenous congestion (see the discussion in section **??** and Appendix A.6 for details of the empirical estimation). In both the exogenous case and the endogenous case, the subsidy increases overall ridership. The change in overall ridership, measured by $\frac{dx_1}{ds} + \frac{dx_2}{ds}$, is 0.047 trips per day in the exogenous case and 0.046 trips per day in the endogenous case. Thus, the endogenous case has a two percent lower increase in overall ridership than the exogenous case.

Table 4 also suggests that MSC_1 and MSC_2 are both less than the price in the base case. This means that there is under-consumption of rides relative to the first best in both periods (where price would be set equal to the marginal social cost). This relationship helps explain some of our welfare findings. We also consider a case for a crowded corridor where MSC_2 is less than the price.[29]

5.2 Natural Experiment: Welfare estimates

5.2.1 Estimates for the Experimental Sample

We consider the welfare impacts of two subsidies in detail: one is a general subsidy for all riders in the experiment 1, and the second is a targeted subsidy for riders on what we call the Transbay route. We define the Transbay route to include trains that lead to the Transbay Corridor[30], and also trains that travel on the Transbay route westbound. This route is in the top 10 percent of congested routes for BART in the morning.[31] We conclude by briefly considering the welfare effects of an optimal subsidy.

Table 5 presents the main welfare results for the base case subsidy and a targeted subsidy for the Transbay route. First, we consider the base case. We measure welfare in dollars per person per weekday. Because the price is above the marginal social cost in both periods in the base case, an increase (decrease) in ridership in a particular period will increase (decrease) welfare in that period (see equation (1)). Ridership decreases in the peak period as a result of

[29]The average fare exceeding the marginal social cost may be unusual in rapid transit systems. For example, Parry and Small (2009) find that the marginal private cost, and thus the marginal social cost, exceeds price for rail in Washington D.C., Los Angeles and London. Winston and Maheshri (2007) find similar results.

[30]See "Transbay Corridor Core Capacity Program Train Control Modernization Project", BART.

[31]This includes routes that depart or arrive at stations: West Oakland, Embarcadero, Bay Fair, San Leandro, Coliseum, Fruitvale, Lake Merritt, Orinda, Rockridge, MacArthur, 19th St/Oakland, 12th St/Oakland City Centre, and Ashby, towards the Transbay tube direction.

the subsidy, and there is a decline in welfare (see row 1). Ridership increases in the off-peak period, and there is an increase in welfare (see row 2).

Table 5: Welfare impacts for the base case and a targeted subsidy

	Parameter	Base case			Targeted subsidy		
		Exogenous congestion	Endogenous congestion	Percent Change	Exogenous congestion	Endogenous congestion	Percent Change
(1)	Welfare change associated with peak travel change	-0.0037	-0.0032	-14.7	0.0006	0.0005	-14.7
(2)	Welfare change associated with off-peak travel change	0.0328	0.0309	-5.9	0.0328	0.0309	-5.9
(3)	Total Welfare change	0.0291	0.0277	-4.8	0.0334	0.0314	-6.1

Notes: Welfare is measured in dollars per person per weekday. Overall welfare increases with the subsidy. The decline in ridership in the peak period reduces welfare, and the increase in ridership in the off-peak period increases welfare. See text for details. The total welfare change is calculated from equation (1).

The overall welfare effect for the base case is the combination of the welfare loss associated with reduced peak travel and the welfare gain associated with more off-peak travel. For both the exogenous and the endogenous congestion cases, this net gain is positive. A subsidy of \$0.25 per trip in the off-peak period results in a net welfare gain of about \$0.03 per rider per day for the exogenous congestion case (row 3 of Table 5).

Welfare in the endogenous case declines by 4.8% compared with the exogenous case because of differences in demand response (see row 3 of the table). This percentage change masks more significant differences in actual travel patterns in the peak and off-peak periods (see the percentage change in row 1 and row 2 for the base case in Table 5).

The welfare analysis for the Transbay route is similar to the base case qualitatively. The main point of this analysis is to illustrate that there could be potential gains from targeting. The average net benefits are about 15% higher when targeting Transbay riders under the assumption of exogenous congestion.[32]

Average net benefits per dollar of direct subsidy, which sometimes is used as a measure of welfare by decision makers, is also higher.[33] This measure ignores the fiscal externality, which is discussed below in estimating the MVPF. We estimate average net benefits by taking the total welfare change and dividing by the average subsidy expenditure per person per weekday.[34]

[32]Take the average welfare gain of .0334 for Transbay and divide by .0291 for the base case. Note that this percentage is higher in the endogenous case.

[33]The measure is similar to that used in the tax literature that estimates the welfare change per dollar of tax revenue. See, *e.g.*, Saez, Slemrod and Giertz (2012).

[34]See appendix A.10 for details.

For the base case, we find that an essential difference between the analysis of the Transbay route and our analysis of the base case is the level of congestion, which is higher. This, in turn, increases the MSC associated with the Transbay route during the peak period. The average peak period congestion on the Transbay route is 108 passengers per car, while it is 57 passengers per car for all other routes.[35] We also find that delay is more severe on the Transbay route than other routes. Using equation (1), we obtain a welfare estimate of providing a subsidy that exclusively focuses on the Transbay route during the morning (see appendix A.8 for details of the calculation).

For the Transbay route, the marginal social cost per peak period trip is \$4.12. This is higher than the marginal social cost when we use the average level of congestion across all routes (\$3.22). The welfare effect of a \$0.25 subsidy for an off-peak trip is \$0.033 per person per day in the case of exogenous congestion, and \$0.031 in the case of endogenous congestion. These welfare estimates are greater than the welfare effect effects estimated in the base case. The difference is due to the higher marginal social cost in the peak period for the Transbay route.

With additional assumptions about the functional form of demand, we can compute the optimal subsidy using equation (A5).[36] We provide details of this computation in appendix A7. The analysis relies on the assumption that demand is linear over a wide range, which may not be the case. With this caveat in mind, we find that the optimal subsidy is about eight times higher than the actual subsidy (\$1.87 in the exogenous case and \$1.90 in the endogenous case versus an actual subsidy of \$0.25 per trip). Total welfare is about the same for the exogenous and endogenous cases, but is about four times higher than welfare associated with the actual subsidy (\$.12 vs. \$.03). The reason that total welfare is similar in the endogenous and exogenous cases is because the optimal subsidy is slightly higher in the endogenous case, but the overall demand increase ($\frac{dx_1}{ds} + \frac{dx_2}{ds}$) is slightly lower. These changes tend to cancel each other out (see equation 1) (see appendix A7 for details).

We also considered what would happen with changes in the shape of the MEC function. In particular, we allow for the MEC function for peak period travel to be an increasing linear and non-linear function of density, e, instead of a constant function of e. Our main conclusion is that in this particular example, the results do no change much in moving from a constant MEC function to a linear MEC function or in moving from a constant MEC function to a quadratic MEC function for the peak period. For example, the total welfare change for the exogenous case is about .029 per person per day with a constant MEC function, .029 for a

[35] The units of the congestion measure used in our calculation is passengers per square meter in an average car. See appendix A.7 for details.

[36] We consider the optimal subsidy for the off-peak market. This is second-best because prices do not equal the marginal social cost in both markets. We could in principle compute optimal subsidies in both markets. However, to compute the welfare change, we would also need an estimate of how demand changes with price in the peak period, which we do not have.

linear MEC function and .028 for a quadratic MEC function. The results are also similar for endogenous congestion (.028, .028 and .027). Overall welfare decreases between exogenous and endogenous congestion are also similar, on the order of 5%. We discuss our methodology and present the empirical results in appendix A.12.

The preceding analysis of welfare changes in Table 5 does not explicitly consider how changes in other modes of travel could affect our calculation. In Appendix A.10, we present a calculation of how intermodal substitution could affect the welfare from the subsidy. In general, we argue that welfare would improve because of decreases in emissions and reduced congestion that are not priced in the market, supporting previous research by Anderson (2014). As an illustrative example, we assume that the overall miles traveled remains constant, but that the increase in the use of BART as a result of the subsidy is accounted for by a reduction in car travel. We compute the benefits associated with these improvements for the case in which the subsidy includes all riders, and all riders are assumed to respond similarly to those in the experiment. The estimated benefits include local pollution reduction benefits of \$0.07 million per year and global pollution benefits associated with CO_2 reductions of \$0.25 million per year, for a total of about \$0.32 million per year. These pollution reduction benefits may be compared with the benefits of the subsidy from the experiment, which are about \$2 million per year. We derive this estimate by multiplying the average benefit per person per weekday by the number of weekdays per year by the number of BART riders. See Appendix A.10 for details of the calculation.[37].

5.2.2 Estimates for the BART Network

We extend our theoretical framework to allow us to estimate the welfare benefits when riders may differ in their responsiveness to a subsidy. We do this to develop a more realistic welfare estimates when data are available.

Formally, we model two types of riders with different demand characteristics. We allow for differences in demand characteristics because people participating in the experiment may, for example, be more sensitive to price changes on average than those who do not. In appendix A.17, we show that this can lead to a straightforward modification of equation 1, where the average per capita welfare change for all riders depends on the demand characteristics of both groups and their respective sizes. For example, when demand characteristics are the same for the two groups, average per capita welfare benefits from the subsidy are the same.

To empirically implement this framework, we first explored how the sample in the natural

[37] We do not analyze how the introduction of the BART network affects pollution. Previous research suggests this could be important (Gendron-Carrier et al., 2022).

experiment differed from the general BART population. As discussed below, the sample does not appear to be that different. Nonetheless, because there may be unobserved differences between populations, we present a sensitivity analysis to assess how welfare could change if the sample under study was not representative of the general population.

A key question is how participants in the natural experiment may differ from the general population that uses BART. We relied on a survey administered by BART to address this question. Our main finding is that, based on the survey, the observable characteristics of our sample and the BART customer sample are no different. For instance, the percent of individuals who have household income less than $50,000 in our natural experiment sample was 13%, versus 14% for all those using BART who commute downtown. The percent of individuals who are BIPOC in the natural experiment sample was 57%, versus 64% for all those using BART who commute downtown. Thus, in terms of external validity on observables, the sample is not that different from all those using BART.

However, there may be unobserved demand differences between those who signed up to the natural experiment and the rest of the BART network. To address this issue, we compared our base case results with a case in which the rest of the BART network is assumed to be one-half as responsive to the price subsidy as participant in the experiment. In the base case, recall the that average net benefit per dollar of direct subsidy was $0.36. If the demand by those not in the experiment is one-half that of people in the experiment, this measure drops by about 46% to $0.2. The reason for the large drop is that other people not in the experiment represent about 93.5 percent of the entire network, and thus have are the major factor driving the per capita welfare calculation. This analysis highlights the need to understand the demand characteristics of different groups of riders to develop a more precise estimate of the welfare impacts of a large-scale subsidy.

5.3 Natural Experiment: MVPF estimates

The preceding section used a measure of welfare based on our welfare model. It assumes that the method of financing the subsidy is a lump sum tax. Net benefits are computed in terms of dollars per person per week day. In contrast, the MVPF calculation for the subsidy is silent on the source of the revenues that fund the intervention.[38]

The MVPF benefits measure is given in terms of the change in net benefits per net additional dollar of government cost. It is, thus, not a welfare measure in the sense of providing an estimate of net benefits (*i.e.*, benefit minus costs). However, if we assume that the net costs to

[38]This source can be considered in a separate MVPF – for example, the MVPF of raising a dollar of funds through an income tax on a particular group.

the government are funded by a lump sum tax, it is possible to show that the MVPF measure and the measure for welfare we obtain in our basic model are related.[39] The intuition is that a lump sum tax converts the net cost to the government into a cost to society.

Equation 8 describes our MVPF calculation.

$$MVPF = \frac{\text{Inframarginal benefit from the subsidy + Change in congestion benefits}}{\text{Change in direct subsidy cost + Change in the transfer to BART}} \tag{8}$$

In our application, the benefit to BART users is the numerator and the net cost to the government of the subsidy is the denominator. The numerator consists of two terms: inframarginal benefits of the subsidy for those off-peak users who did not switch from the peak, and the benefits to all peak travelers from reduced congestion.[40] The denominator also consists of two terms: the direct subsidy costs that are paid to BART consumers, and the change in the transfer to BART resulting from the change in ridership.

The two terms in the numerator are positive. The subsidy provides benefits to those who do not change their usage patterns and there are benefits (in terms of reduced congestion) for those who continue to travel during the peak period. The two terms in the denominator have different signs. The direct subsidy cost to the government increases as a result of introducing the subsidy; however, in our particular example, discussed in more detail below, BART's revenues profits actually increase as a result of the subsidy, which means that the transfer from the government to BART, which is needed to cover BART's losses actually decreases. Thus, the second expression in the numerator is negative. This situation arises because ridership increases as a result of the intervention, and price exceeds marginal cost for each ride. BART's profits thus increase from the increased ridership, and the transfer needed by the government to keep BART's profits at a particular level decreases.

Table 6 shows the MVPF calculation for the cases of exogenous and endogenous congestion. In both cases the MVPF is close to 1.6. This means that for each additional dollar of net cost to the government, the marginal benefit to all parties is \$1.60. The 95% confidence interval for the MVPF, obtained by bootstrapping both the peak and off-peak demand response estimations, ranges from about 1.4 to 1.7 for both cases.[41]

There are two general points to be made about the MVPF approach based on our example.

[39] They are related in the sense that the welfare measure $\frac{dW(s)}{ds}$ is the same using both approaches under this assumption. See appendix A.15 for details. A key feature common to both approaches is their reliance on the envelope theorem.

[40] We assume the net benefits to those who shift from peak to off-peak travel as a result of the subsidy are negligible or zero because of the envelope theorem, and do not include this term in net benefits. If we included some part of these benefits, it would have the effect of increasing the MVPF.

[41] If we assume riders not in the experiment were half as sensitive to the subsidy as those in the experiment then the MVPF drops from 1.6 to 1.26.

Table 6: MVPF's for the subsidy

	Exogenous Congestion	Endogenous Congestion
Numerator		
1. Inframarginal benefits	0.307	0.307
2. Congestion benefits	0.0258	0.0220
Numerator total	0.3328	0.329
Denominator		
1. Change in direct subsidy cost	0.307	0.307
2. "Change in transfer to BART"	-0.0987	-0.0965
Denominator Total	0.2083	0.211
MVPF = Numerator/Denominator and error bounds	1.598 [1.479, 1.716]	1.563 [1.444, 1.681]

Notes: The MVPF is computed for a subsidy of $1 per trip in the off-peak period using equation (8). The VOT is set equal to 50% of the wage rate. The 95% confidence interval is calculated by bootstrapping the estimation for the demand responses with 100 replications. Details of the calculation are provided in appendix A.7.

First, the subsidy in this case involves more than simply the direct subsidy to passengers. Because we are dealing with a regulated entity, one must also consider the impact of the subsidy on the revenues needed to support the regulated entity – in this case, BART. This is a feature that is endemic to many regulatory activities, where prices depart from marginal cost (e.g., Hahn and Metcalfe (2021)).

Second, more work needs to be done in this area to rank various subsidies and pricing interventions using the MVPF framework in the regulatory domain, and connecting them to tax policies that might be used to assess their effectiveness. That will allow us to make more informed assessments of the actual welfare impacts of policies. A review of Hendren (2020) reveals the MVPFs we estimate are comparable to, and in some cases higher than, many educational and labor policy interventions.[42]

5.4 Field Experiment: MVPF Estimates

This section analyzes the welfare effect of the targeted subsidies used in experiment 2. In the interest of brevity, we focus on results using the MVPF here, and present results on the net

[42] A higher MVPF does not necessarily imply the policy is "better." One would need to make further assumptions about equity weights to compare policies.

welfare benefits in the appendix. The results are qualitatively similar.[43]

There are two main results. First, targeting results in many MVPFs that exceed one (meaning the increase in willingness to pay exceeds the net cost to the government), but in some that do not. This is because targeting parameters are set without reference to changes in demand, and these changes are critical for welfare. Second, we explain the subtle relationship between a change in the value of time and a change in welfare. The key point is that the value of time interacts with the demand changes in both periods to affect the change in welfare. As a result, the impact of a change in the VOT on economic welfare cannot be signed *apriori*.

In experiment 2, thirty groups receive a subsidy in 12 time periods that are 20 minutes long. The subsidies are to encourage travel that is 20 minutes or 40 minutes away from the rider's usual travel period of 20 minutes. A group is defined by their subsidized time period and their usual period of travel. Because of data limitations, we assume that the targeted subsidy for a specific group only affects travel in the targeted period and the usual travel time period.

We identify each targeted group, j, by the usual travel interval and the subsidized interval, which can be 20 or 40 minutes before or after the usual travel time. For example, one group is riders who usually travel from 8:20 am to 8:40 am and receive a subsidy to travel between 8.40 am and 9 am.

We focus on evaluating the welfare impact of the subsidies for experiment 2 using the MVPF. This is because each of the targeting subsidies involves a small fraction of riders among all BART users. We consider the case in which the riders take congestion as exogenous in their decision making. In this case, the MVPF provides a simple measure that allows us to rank the welfare impact of each of the subsidies.[44]

In calculating the MVPF for subsidies in experiment 2, we allow for the possibility that the marginal external cost is positive in the targeted interval. We calculate the MVPF for each subsidy using the following formula:

$$MVPF_{jt} = \frac{x_{jt} - MEC_{j1} * \frac{dx_{j1}}{ds_{jt}} - MEC_{jt} * \frac{dx_{jt}}{ds_{jt}}}{x_{jt} - (p - MPC) * (\frac{dx_{j1}}{ds_{jt}} + \frac{dx_{jt}}{ds_{jt}})} \tag{9}$$

$MVPF_{jt}$ is the MVPF of the subsidy given to group j for the targeted time interval t; p is the price for a trip; MEC_{j1} is the marginal external cost for the route taken by group j on the usual time before the subsidy, and MEC_{jt} is the marginal external cost for the same route in the time targeted by the subsidy. x_{jt} is the average number of trips taken by group j in the

[43] See appendix B Figure A16.

[44] Our qualitative findings are the same when we use our welfare model in section A.1 to estimate welfare changes.

targeted travel time. The numerator consists of two parts: the inframarginal benefits for those who receive the subsidy in time t (the first term) plus the change in external congestion costs (the last two terms). The denominator consists of two parts as well: the direct costs to the government of the subsidy for customers (the first term) plus the change in profits to BART (the second term).

We measure the congestion level using the passenger density on the route travelled by group j, at the usual travel time and also at the subsidized time. We also consider the travel delay in estimating the marginal external cost in each time period. This allows us to calculate MEC_{j1} and MEC_{jt}.

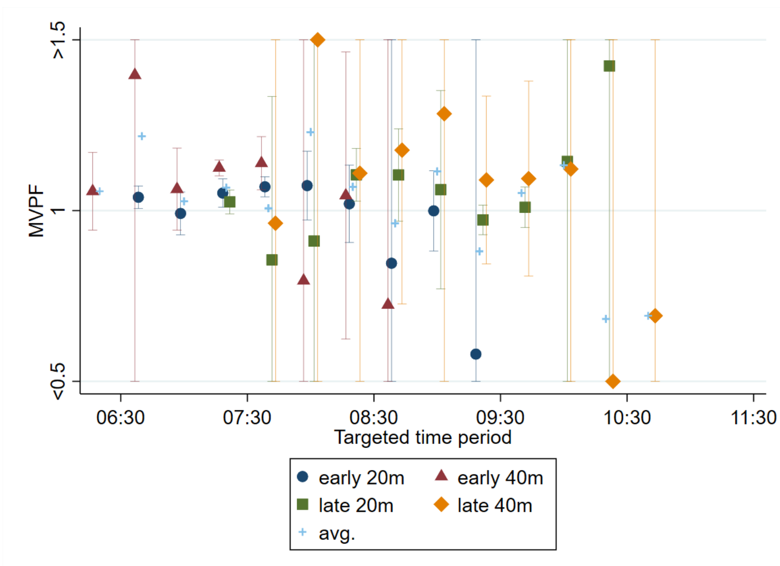
To calculate the marginal external cost, we need a measure of the value of time (see equation (A9)). Using survey data from the sample, the median wage for a single person household is \$50 dollar per hour. We focus on these households because the survey did not ask for the number of working individuals in a household, but did ask for income and the number of people in the household. This wage estimate is higher than the median wage used to evaluate experiment 1, which was based on all San Francisco residents. We consider a value of time equal to 50% of the wage, as we did before, and also one equal to 75% of the wage. This allows us to calculate the marginal external and social cost using equation (A9), as we did for experiment 1.

We use the estimate of marginal private cost per trip discussed in section 2. We assume the price of an average trip is about \$4 for all groups. The behavioral parameters, $\frac{dx_{j1}}{ds_{jt}}$ and $\frac{dx_{jt}}{ds_{jt}}$, are based on the analysis in section 3. Table A11 in section A.13 presents the parameter values we used. Figure 7 plots the MVPFs for each of the subsidies in experiment 2, by group. The figure shows that the point estimates of the MVPFs of these targeted subsidies can be above or less than 1 for both the case where VOT is 50% of the wage rate (panel a) (Small, 2012) and 75% of the wage rate (panel b) (Goldszmidt et al., 2020). For example, the MVPF for the subsidy that shifts riders who usually travel in 8:00-8:20am to travel during 8.40-9.00am exceeds 1; the MVPF for the subsidy that shifts riders who usually travel from 9:20-9:40am to travel during 8:40-9:00am is less than 1. The average MVPF for a value of time at 50% of the wage rate is 1, but for a value of time at 75% of the wage rate the average MVPF is 1.17.

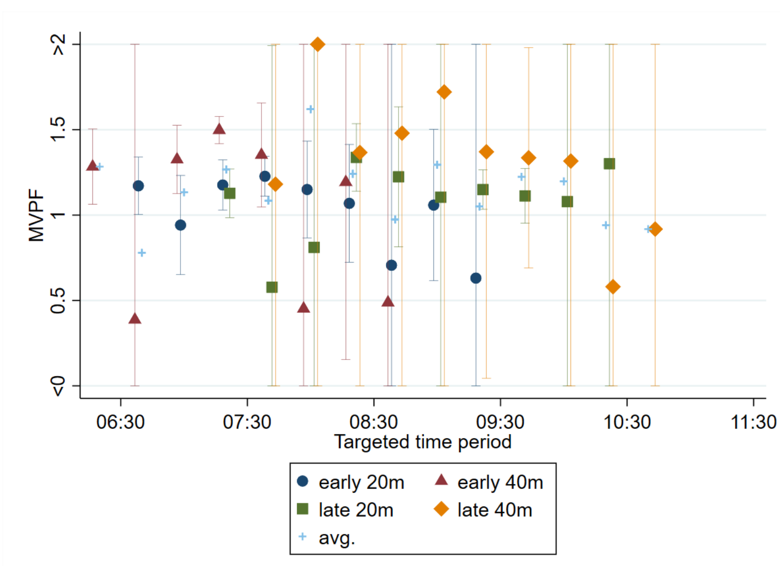
The impact of a change in the value of time cannot be signed apriori. A change in the value of time enters into the analysis through its impact on the MEC in both periods. If the change in the value of time entered into the calculation in only one period, the comparative statics would be straightforward. However, when we allow for a change in the MEC in both periods, the comparative statics depends on the MEC *and* demand parameters in both periods, and this effect cannot be signed a priori. We find that in 75.7 percent of cases, an increase in value of time from 50 to 75 percent increases welfare, and in 24.3 percent of cases an increase in

Figure 7: MVPF of subsidy in experiment 2

(a) VOT - 50% of wages



(b) VOT - 75% of wages



Note: This figure shows the MVPF of a targeted subsidy. See text for details. The x-axis represents the beginning time of the targeted time period. The shape of each point shows the usual travel time of the group of riders, which is 20 or 40 minutes earlier or later than the targeted time. The value of time is assumed to be 50% of the wage in panel (a) and 75% of the wage in panel (b).

value of time decreases the welfare effect. [45]

A comparison of the MVPFs from the two experiments suggests that the average MVPF from the field experiment is lower than the MVPF obtained in the natural experiment. This may at first seem unusual because one might expect net benefits to be higher in the case of targeting (experiment 2) than with a general off-peak subsidy (experiment 1). However, there are several possible reasons for this finding. First, the MVPF provides a measure of marginal and not total net benefits. A social planner could in principle use the information from a finely targeted experiment like the field experiment to achieve higher total net benefits than an experiment that simply provided the same subsidy to all users (e.g, experiment 1). This increase in net benefits could in principle be achieved with differentiated subsidies across different time intervals. Second, the two experiments are measuring different things. The natural experiment measures the marginal benefit of moving people from peak to off-peak, and the field experiment measures the marginal benefit of moving from one relatively narrow time interval to an adjacent time interval, which is not necessarily off-peak. Finally, the estimate of congestion used in the field experiment varies across all of the relatively small time intervals, whereas we assume for simplicity in our natural experiment that there is one average level of congestion that applies to the peak period, and there is no congestion in the off-peak period. The bottom line is that there is no reason to assume a particular relationship between the MVPFs estimated in the two experiments, but the more finely tuned approaches could lead to higher net benefits from targeting.

6 Conclusion

Using a large natural experiment and a natural field experiment, we develop novel estimates of peak and off-peak price elasticities for urban mass transit demand in San Francisco. Both experiments allow us to obtain causal estimates of demand.

Using the demand estimates, we assess the welfare impacts of the price subsidies using a sufficient statistics approach. Our analysis suggests that off-peak subsidies can increase welfare, but the positive effects are reduced when consumers take the decisions of others into account compared to when they do not. We also find a large variation in the welfare impacts of shifting travel to different periods, which is explained by differences in demand and congestion

[45] The precise formula for measuring the impact of VOT on MVPF is obtained by differentiating equation (9) with respect to value of time. It is given by $-(\frac{dMEC_{j1}}{dVOT}\frac{dx_{j1}}{ds_{jt}}+\frac{dMEC_{jt}}{dVOT}\frac{dx_{jt}}{ds_{jt}})MVPF_{jt}$, where VOT represent value of time, and $\frac{dMEC_{j1}}{dVOT}$ and $\frac{dMEC_{jt}}{dVOT}$ are the change in each of the MEC with respect to VOT, that depends on the crowding level during the time period, and how the delay relates to the number of passengers. See appendix A.4 for the derivation and further discussion of the empirical results.

characteristics. Finally, we show that the targeting of subsidies can increase welfare, but need not do so if the regulator does not have accurate information on demand.

The experiments offer different insights into the impact of providing subsidies at particular times for economic welfare. We find that the subsidy in the natural experiment generally increases welfare. In the base case, we show that the marginal value of public funds is 1.6. We also calculate the average net benefits per dollar of subsidy in the natural experiment, and find that the average net benefits are $0.36 per dollar of subsidy expenditure. In contrast, for the field experiment, MVPFs showed considerably more variation depending on the particular time period being targeted.

There are three areas for future research that we believe would be useful. First, with respect to the issue of managing transport networks, we think it would be instructive to explore the extent to which various forms of congestion pricing result in attractive MVPFs. In particular, we would advocate doing more experiments that allow for causal identification of key elasticities related to peak and off-peak pricing. This would allow scholars to draw more robust conclusions about the generalizability of studies in this area (List, 2020, 2022). In addition, we think it would be useful to have experiments that focused on both the intensive margin (*e.g.*, making greater use of a transport mode that a person already uses) and the extensive margin (*e.g.*, changing transport modes or changing the city structure (Desmet and Rossi-Hansberg, 2013; Severen, 2019; Tsivanidis, 2019; Heblich, Redding and Sturm, 2020; Barwick et al., 2021)).

Second, as pointed out by Kleven (2021), sufficient statistics is only one of many approaches for estimating the impacts of economics policies. A strength of this approach is its use of the envelope theorem in simplifying the analysis. But this may also be a weakness when dealing with non-marginal policy changes. While non-marginal changes can, in principle, be estimated using our welfare model and the MVPF approach (Hendren, 2020), it is an empirical question as to whether such estimates are reasonably accurate. It would be instructive to explore how different models yield different results for a particular class of problems, such as the one studied here. It would also be useful to explore the relationship between sufficient statistics approaches designed to measure welfare, such as those developed by Chetty (2009), and the MVPF approach advanced by Hendren (2020).

Third, we noted in our analysis that the targeting of subsidies does not always meaningfully increase welfare when the decision maker does not have perfect information. We think that empirical and theoretical work could help identify when targeting actually helps.

References

Aiken, Emily, Suzanne Bellue, Dean Karlan, Chris Udry, and Joshua E Blumenstock. 2022. "Machine learning and phone data can improve targeting of humanitarian aid." *Nature*, 603(7903): 864–870.

Alatas, Vivi, Ririn Purnamasari, Matthew Wai-Poi, Abhijit Banerjee, Benjamin A Olken, and Rema Hanna. 2016. "Self-targeting: Evidence from a field experiment in Indonesia." *Journal of Political Economy*, 124(2): 371–427.

Allcott, Hunt, and Dmitry Taubinsky. 2015. "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market." *American Economic Review*, 105(8): 2501–38.

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky. 2019. "Regressive sin taxes, with an application to the optimal soda tax." *Quarterly Journal of Economics*, 134(3): 1557–1626.

Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky. 2014. "Energy policy with externalities and internalities." *Journal of Public Economics*, 112: 72–88.

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

Athey, Susan. 2018. "The impact of machine learning on economics." In *The economics of artificial intelligence: An agenda*. 507–547. University of Chicago Press.

Barwick, Panle Jia, Shanjun Li, Andrew R Waxman, Jing Wu, and Tianli Xia. 2021. "Efficiency and equity impacts of urban transportation policies with equilibrium sorting." National Bureau of Economic Research.

Boiteux, Marcel. 1960. "Peak-load pricing." *The Journal of Business*, 33(2): 157–179.

Borenstein, Severin. 2005. "The long-run efficiency of real-time electricity pricing." *Energy Journal*, 26(3).

Bundorf, M Kate, Jonathan Levin, and Neale Mahoney. 2012. "Pricing and welfare in health plan choice." *American Economic Review*, 102(7): 3214–48.

Burkhardt, Jesse, Kenneth Gillingham, and Praveen K Kopalle. 2019. "Experimental evidence on the effect of information and pricing on residential electricity consumption." National Bureau of Economic Research.

Burlig, Fiona, Christopher Knittel, David Rapson, Mar Reguant, and Catherine Wolfram. 2020. "Machine learning from schools about energy efficiency." *Journal of the Association of Environmental and Resource Economists*, 7(6): 1181–1217.

Byrne, David P, Leslie A Martin, and Jia Sheen Nah. 2019. "Price Discrimination, Search, and Negotiation in an Oligopoly: A Field Experiment in Retail Electricity."

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

Chassang, Sylvain, Padró I Miquel, Erik Snowberg, et al. 2012. "Selective trials: A principal-agent approach to randomized controlled experiments." *American Economic Review*, 102(4): 1279–1309.

Chetty, Raj. 2009. "Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods." *Annual Review of Economics*, 1(1): 451–488.

Christensen, Peter, Paul Francisco, Erica Myers, Hansen Shao, and Mateus Souza. 2021. "Energy Efficiency Can Deliver for Climate Policy: Evidence from Machine Learning-Based Targeting."

Cohen, Jessica, Pascaline Dupas, and Simone Schaner. 2015. "Price subsidies, diagnostic tests, and targeting of malaria treatment: evidence from a randomized controlled trial." *American Economic Review*, 105(2): 609–45.

Davis, Lucas W. 2021. "Estimating the price elasticity of demand for subways: Evidence from Mexico." *Regional Science and Urban Economics*, 87: 103651.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review*, 110(9): 2964–96.

Desmet, Klaus, and Esteban Rossi-Hansberg. 2013. "Urban accounting and welfare." *American Economic Review*, 103(6): 2296–2327.

Domencich, Thomas A, and Daniel McFadden. 1975. "Urban travel demand-a behavioral analysis."

Dupas, Pascaline. 2014. "Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment." *Econometrica*, 82(1): 197–228.

Farhi, Emmanuel, and Xavier Gabaix. 2020. "Optimal taxation with behavioral agents." *American Economic Review*, 110(1): 298–336.

Finkelstein, Amy, and Nathaniel Hendren. 2020. "Welfare analysis meets causal inference." *Journal of Economic Perspectives*, 34(4): 146–67.

Fowlie, Meredith, Catherine Wolfram, Patrick Baylis, C Anna Spurlock, Annika Todd-Blick, and Peter Cappers. 2021. "Default effects and follow-on behaviour: Evidence from an electricity pricing program." *The Review of Economic Studies*, 88(6): 2886–2934.

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew A Turner. 2022. "Subways and urban air pollution." *American Economic Journal: Applied Economics*, 14(1): 164–96.

Gerarden, Todd D, and Muxi Yang. 2021. "Using Targeting to Optimize Program Design: Evidence from an Energy Conservation Experiment." Tech. rep., Working Paper.

Glaister, Stephen. 1974. "Generalised consumer surplus and public transport pricing." *Economic Journal*, 84(336): 849–867.

Glaister, Stephen, and Davis Lewis. 1978. "An integrated fares policy for transport in London." *Journal of Public Economics*, 9(3): 341–355.

Goldszmidt, Ariel, John A List, Robert D Metcalfe, Ian Muir, V Kerry Smith, and Jenny Wang. 2020. "The Value of Time in the United States: Estimates from Nationwide Natural Field Experiments." National Bureau of Economic Research.

Hahn, Robert W, and Robert D Metcalfe. 2021. "Efficiency and Equity Impacts of Energy Subsidies." *American Economic Review*, 111(5): 1658–88.

Harberger, Arnold C. 1964. "The measurement of waste." *American Economic Review*, 54(3): 58–76.

Harrison, Glenn W, and John A List. 2004. "Field experiments." *Journal of Economic literature*, 42(4): 1009–1055.

Haywood, Luke, and Martin Koning. 2015. "The distribution of crowding costs in public transport: New evidence from Paris." *Transportation Research Part A: Policy and Practice*, 77: 182–201.

Heblich, Stephan, Stephen J Redding, and Daniel M Sturm. 2020. "The making of the modern metropolis: evidence from London." *Quarterly Journal of Economics*, 135(4): 2059–2133.

Hendren, Nathaniel. 2013. "Private information and insurance rejections." *Econometrica*, 81(5): 1713–1762.

Hendren, Nathaniel. 2020. "Measuring economic efficiency using inverse-optimum weights." *Journal of Public Economics*, 187: 104198.

Holmgren, Johan. 2007. "Meta-analysis of public transport demand." *Transportation Research Part A: Policy and Practice*, 41(10): 1021–1035.

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2018. "Moral suasion and economic incentives: Field experimental evidence from energy demand." *American Economic Journal: Economic Policy*, 10(1): 240–67.

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2021. "Selection on welfare gains: Experimental evidence from electricity plan choice." National Bureau of Economic Research.

Jacobsen, Mark R, Christopher R Knittel, James M Sallee, and Arthur A Van Benthem. 2020. "The use of regression statistics to analyze imperfect pricing policies." *Journal of Political Economy*, 128(5): 1826–1876.

Jacobsen, Mark R, James M Sallee, Joseph S Shapiro, and Arthur A van Benthem. 2022. "Regulating Untaxable Externalities: Are Vehicle Air Pollution Standards Effective and Efficient?"

Jansson, Kjell. 1993. "Optimal public transport price and service frequency." *Journal of Transport Economics and Policy*, 33–50.

Jessoe, Katrina, and David Rapson. 2014. "Knowledge is (less) power: Experimental evidence from residential energy use." *American Economic Review*, 104(4): 1417–38.

Joskow, Paul L, and Catherine D Wolfram. 2012. "Dynamic pricing of electricity." *American Economic Review*, 102(3): 381–85.

Kleven, Henrik J. 2021. "Sufficient Statistics Revisited." *Annual Review of Economics*, 13: 515–538.

Knittel, Christopher R, and Ryan Sandler. 2013. "The welfare impact of indirect pigouvian taxation: Evidence from transportation." National Bureau of Economic Research.

Knittel, Christopher R, and Samuel Stolper. 2019. "Using machine learning to target treatment: The case of household energy use." National Bureau of Economic Research.

Kreindler, Gabriel. 2022. "Peak-hour road congestion pricing: Experimental evidence and equilibrium implications." *Unpublished paper*.

Larcom, Shaun, Ferdinand Rauch, and Tim Willems. 2017. "The benefits of forced experimentation: striking evidence from the London underground network." *Quarterly Journal of Economics*, 132(4): 2019–2055.

Levin, Laurence, Matthew S Lewis, and Frank A Wolak. 2017. "High frequency evidence on the demand for gasoline." *American Economic Journal: Economic Policy*, 9(3): 314–47.

List, John A. 2020. "Non est disputandum de generalizability? A glimpse into the external validity trial." National Bureau of Economic Research.

List, John A. 2022. *The voltage effect: How to make good ideas great and great ideas scale.* Currency.

Litman, Todd. 2004. "Transit price elasticities and cross-elasticities." *Journal of Public Transportation*, 7(2): 3.

Mahoney, Neale, and E Glen Weyl. 2017. "Imperfect competition in selection markets." *Review of Economics and Statistics*, 99(4): 637–651.

McFadden, Daniel. 1974. "The measurement of urban travel demand." *Journal of Public Economics*, 3(4): 303–328.

McFadden, Daniel. 2001. "Economic choices." *American Economic Review*, 91(3): 351–378.

McFadden, Daniel, Antti Talvitie, Stephen Cosslett, Ibrahim Hasan, Michael Johnson, Fred Reid, and Kenneth Train. 1977. "Demand model estimation and validation." *Urban Travel Demand Forecasting Project, Phase*, 1.

McFadden, Daniel, William B Tye, and Kenneth Train. 1977. *An application of diagnostic tests for the independence from irrelevant alternatives property of the multinomial logit model.* Institute of Transportation Studies, University of California Berkeley, CA.

Mohring, Herbert. 1972. "Optimization and scale economies in urban bus transportation." *The American Economic Review*, 62(4): 591–604.

Mullainathan, Sendhil, and Jann Spiess. 2017. "Machine learning: an applied econometric approach." *Journal of Economic Perspectives*, 31(2): 87–106.

Oum, Tae Hoon, and Michael W Tretheway. 1988. "Ramsey pricing in the presence of externality costs." *Journal of Transport Economics and Policy*, 307–317.

Parry, Ian WH, and Kenneth A Small. 2009. "Should urban transit subsidies be reduced?" *American Economic Review*, 99(3): 700–724.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva. 2014. "Optimal taxation of top labor incomes: A tale of three elasticities." *American Economic Journal: Economic Policy*, 6(1): 230–71.

Polyakova, Maria, and Stephen P Ryan. 2019. "Subsidy targeting with market power." National Bureau of Economic Research.

Reguant, Mar. 2019. "The efficiency and sectoral distributional impacts of large-scale renewable energy policies." *Journal of the Association of Environmental and Resource Economists*, 6(S1): S129–S168.

Rodrik, Dani. 1987. "Policy targeting with endogenous distortions: Theory of optimum subsidy revisited." *Quarterly Journal of Economics*, 102(4): 903–911.

Saez, Emmanuel, Joel Slemrod, and Seth H Giertz. 2012. "The elasticity of taxable income with respect to marginal tax rates: A critical review." *Journal of Economic Literature*, 50(1): 3–50.

Severen, Christopher. 2019. "Commuting, labor, and housing market effects of mass transportation: Welfare and identification." *Review of Economics and Statistics*, 1–99.

Small, Kenneth A. 2012. "Valuation of travel time." *Economics of transportation*, 1(1-2): 2–14.

Spulber, Daniel F. 1989. *Regulation and markets.* MIT press.

Tarduno, Matthew. 2022. "For Whom the Bridge Tolls: Congestion, Air Pollution, and Second-Best Road Pricing." *Unpublished manuscript*.

Tizghadam, Ali, Hamzeh Khazaei, Mohammad HY Moghaddam, and Yasser Hassan. 2019. "Machine learning in transportation."

Train, Kenneth. 1977. "Optimal transit prices under increasing returns to scale and a loss constraint." *Journal of Transport Economics and Policy*, 185–194.

Tsivanidis, Nick. 2019. "Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio." *Unpublished manuscript*.

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

Weyl, E Glen, and Michal Fabinger. 2013. "Pass-through as an economic tool: Principles of incidence under imperfect competition." *Journal of Political Economy*, 121(3): 528–583.

Williams, Kevin R. 2021. "The welfare effects of dynamic pricing: Evidence from airline markets." National Bureau of Economic Research.

Winston, Clifford. 1985. "Conceptual developments in the economics of transportation: an interpretive survey." *Journal of Economic Literature*, 23(1): 57–94.

Winston, Clifford. 2013. "On the performance of the US transportation system: Caution ahead." *Journal of Economic Literature*, 51(3): 773–824.

Winston, Clifford, and Vikram Maheshri. 2007. "On the social desirability of urban rail transit systems." *Journal of Urban Economics*, 62(2): 362–382.

Wolak, Frank A. 2007. "Residential customer response to real-time pricing: The Anaheim critical peak pricing experiment."

Yang, Nan, and Yong Long Lim. 2018. "Temporary incentives change daily routines: Evidence from a field experiment on Singapore's subways." *Management Science*, 64(7): 3365–3379.

A APPENDIX

This appendix follows the presentation in the paper, providing extensions and derivations related to the theory, empirical analysis of demand, and the empirical analysis of welfare. It consists of 17 sections noted below.

of these subsidies.

Appendix A.10 - Additional discussion for welfare analysis

This section discusses how we include the benefits from reduced externalities from other modes of transit for the calculation of our welfare estimates for the natural field experiment.

Appendix A.11 - Additional estimates on welfare impact

This section provides a numerical example for how we calculate the annual welfare benefits from the subsidy.

Appendix A.12 - Natural experiment welfare estimates when the marginal social cost varies

This section compares the impact of changing the marginal social cost with a change in peak travel with the case when marginal social cost is held constant.

Appendix A.13 - Parameters for welfare analysis of the field experiment

This section presents the parameters used for the welfare analysis of the field experiment, including demand estimates, marginal external cost, and marginal private cost.

Appendix A.14 - Welfare estimates for the natural field experiment using the welfare model

This section presents the welfare estimates for a targeted subsidy for the field experiment using our welfare model.

Appendix A.15 - Comparing the Welfare Model with the MVPF calculation

This section explains how our welfare model relates to the marginal value of public funds approach when the subsidy is funded by a lump sum tax.

Appendix A.16 - Comparing the demand response with exogenous congestion and endogenous congestion

This section shows conditions under which the demand response to a subsidy is smaller in absolute value for the case of endogenous congestion for both the peak and the off-peak periods.

Appendix A.17 - Welfare analysis of Perks subsidy with different demand responses

This section extends our basic theoretical welfare model to allow for two types of riders with different demand responses.

A.1 Welfare model

We develop a welfare model that allows us to estimate the economic welfare effects of a subsidy for off-peak consumers.[46] Consumers maximize utility and the firm minimizes costs subject to meeting demand.

Our key results in this section are intuitive. The welfare effect of a subsidy will be shown to depend on how consumers respond to the subsidy, and the extent to which prices deviate from marginal social cost. If the price is above the marginal social cost in a particular period, and travel increases as a result of the policy intervention (e.g., a subsidy), then it contributes to an increase in welfare. Summing the welfare changes across each period gives the overall change in welfare.

A representative consumer faces the following utility maximization problem:

$$\max_{x_1, x_2, y} u(x_1, x_2, e) + y \qquad \text{s.t.} \quad px_1 + (p-s)x_2 + y = Z$$

Utility, u, is a function of peak-travel (x_1), off-peak travel (x_2), and an externality associated with peak travel (e).[47] Z is income, assumed to be exogenous, and y is a numeraire good with a price of unity that also provides utility. The price per trip for peak travel is p; s is the subsidy given to off-peak riders, and $(p-s)$ is the net price per trip for off-peak travel. This quasi-linear formulation of utility implies there are no income effects.[48]

The utility function has the following standard properties: $u_{x_1} > 0$, $u_{x_2} > 0$, and $u_e < 0$, where these terms denote partial derivatives. We assume that the utility function is strictly quasi-concave in x_1 and x_2, so the solution to the consumer problem is unique. In addition, we assume that the marginal utility of peak travel is weakly decreasing in the level of congestion, i.e., $u_{x_1,e} \leq 0$.

Consumer maximization yields demand functions $x_1(p,s,e)$, $x_2(p,s,e)$ and $y(p,s,e)$. The demand for peak rides, off-peak rides, and the numeraire good all depend on the travel price, the subsidy, and the level of the externality.

We model the regulated firm as one that minimizes costs, subject to the regulator setting prices. In the case of a transport service, such as BART, the problem is typically to minimize

[46] The model could be interpreted as a general equilibrium model, provided the assumptions underlying a sufficient statistics approach apply.

[47] We assume the externality level, e, is determined by a weakly monotonic and differentiable function of total rides in the peak period, so that $e = E(Nx_1)$. This says that the level of the congestion externality is weakly increasing with the total number of peak rides in the system, Nx_1.

[48] The quasi-linear assumption can be relaxed by making different assumptions about the tax system. See Kleven (2021).

current losses, which are defined as operating cost less revenue from fares.

The firm is assumed to be a common carrier that must meet demand for N peak and off-peak riders. In the simplest case, price is set by the regulator at $\bar{p}$. The firm's problem is given by:

$$\pi = \min_p C(Nx_1, Nx_2) - Np(x_1 + x_2)$$

s.t. $p = \bar{p}$, $Nx_1 \leq \bar{x}_1$ and $Nx_2 \leq \bar{x}_2$.

Operating costs are given by the function $C(Nx_1, Nx_2)$, and may include fixed components that do not vary with demand, as well as variable components that do. Fare revenue is given by $pN(x_1 + x_2)$. We assume that the firm knows the consumer demand functions and must meet that demand. The firm's losses are assumed to be financed by a lump sum tax on consumers.

We assume that the firm produces a level of output at the equilibrium, $\bar{x}_1$, in the peak period to meet demand Nx_1; and a level of output, $\bar{x}_2$, in the off-peak period to meet demand Nx_2. Any losses that the firm incurs are assumed to be covered by a lump sum transfer.[49] Furthermore, the firm must charge the same price for peak and off-peak users.

We could make the firm's problem more realistic in several ways. These include allowing for the firm to set price up to some level established by the regulator, as well as considering a game between the firm and the regulator (see, *e.g.*, Spulber (1989)). If, for example, the firm were allowed to set its price subject to a ceiling, then we would need to consider the impact of a subsidy on the firm's price. For relatively small subsidies of the type considered here, the impact of a subsidy on price is likely to be *de minimus*, at least for the current period for which prices are set. We thus opt for a simple formulation of firm behavior.

The welfare function is the sum of consumer utility, firm losses, and the subsidy:

$$W(s) = NV(p,s,e) - \min_p\{C(Nx_1, Nx_2) - Np(x_1 + x_2)\} - sNx_2(p,s)$$

where $V(p,s,e)$ is the indirect utility function of riders at the equilibrium, $C(Nx_1, Nx_2) - Np(x_1 + x_2)$ is the direct transfer to the firm for it to break even, and sNx_2 is the government expenditure for the off-peak subsidy.

The marginal effect of increasing the subsidy is represented by the first derivative of W(s) with respect to s:

$$\frac{dW}{ds} = N\frac{dV}{ds} + \frac{d(-\pi)}{ds} - (\frac{d(-sNx_2)}{ds}) \tag{A1}$$

[49] We add this assumption because it is not uncommon for mass transit systems to receive subsidies.

The first term in the equation is the effect of the subsidy on consumer utility, which equals:

$$\frac{dV}{ds} = x_2 - \frac{\partial u}{\partial e}\frac{\partial e}{\partial x_1}\frac{dx_1}{ds}$$

The indirect effect of the subsidy on consumer utility due to changes in the consumption of x_1 and x_2 (holding congestion constant) is zero because of the envelope theorem. The effect of the subsidy on the consumer's utility thus equals the direct effect of the subsidy, x_2, plus the effect from induced changes in the level of congestion.

The change in BART lost revenues (the second term in equation A1) is

$$\frac{-d\pi}{ds} = N(\frac{dC(Nx_1, Nx_2)}{dNx_1} - p)\frac{dx_1}{ds} + N(\frac{dC(Nx_1, Nx_2)}{dNx_2} - p)\frac{dx_2}{ds}$$

Under the assumption that the regulator sets the price at $p = \bar{p}$, there is no indirect effect from the price change on BART's lost revenues.

The tax revenue change required to finance the subsidy is $\frac{d(-sNx_2)}{ds} = -Nx_2 - N\frac{dx_2}{ds}$.

Summing up terms and dividing by N, the welfare change with respect to a marginal change in the subsidy is:

$$\frac{1}{N}\frac{dW(s)}{ds} = (p - \frac{dC(Nx_1, Nx_2)}{dNx_1})\frac{dx_1}{ds} + (p - \frac{dC(Nx_1, Nx_2)}{dNx_2})\frac{dx_2}{ds} - s\frac{dx_2}{ds} - \frac{\partial u}{\partial e}\frac{\partial e}{\partial x_1}\frac{dx_1}{ds} \quad \text{(A2)}$$

which gives equation (A2).

It is instructive to explain our model in the language of externalities, which generally defines the marginal social cost as the sum of the marginal private cost and the marginal external cost. We refer to MEC_1 as the marginal external cost of congestion for a peak hour trip, which is defined as $-\frac{\partial u}{\partial e}\frac{\partial e}{\partial x_1}$. Furthermore, the change in the level of congestion depends on the change in peak hour travel, which gives the following expression: $\frac{\partial u}{\partial e}\frac{de}{ds} = \frac{\partial u}{\partial e}\frac{\partial e}{\partial x_1}\frac{dx_1}{ds} = -MEC_1\frac{dx_1}{ds}$. We assume that MEC_1 is constant in the interest of simplicity. [50]

We assume the incremental costs of service in the peak period, MPC_1, and the off-peak period, MPC_2, are constant and equal. In particular, $\frac{dC(Nx_1,Nx_1)}{dNx_1} = MPC_1 = \frac{dC(Nx_1,Nx_2)}{dNx_2} = MPC_2$. The marginal social cost of a peak-hour trip, MSC_1, equals the sum of MPC_1 and MEC_1. Since there is no congestion in the off-peak period, the marginal external cost of an off-peak trip is

[50] Appendix A.1 analyzes the case where MEC_1 varies with e. In many applications, including this one, it may be reasonable to assume that MEC is constant over the region of interest, especially if the subsidy is "small". See discussion in section A.12, noting the impact on the change in welfare of allowing MEC to vary is very small (less than 1%).

zero, which implies $MSC_2 = MPC_2$.[51] Both the marginal social cost in the peak period and the marginal social cost in the off-peak period are constant in the basic model.

Equation (A2) can now be rewritten as

$$\frac{1}{N}\frac{dW(s)}{ds} = (p - MSC_1)\frac{dx_1}{ds} + (p - s - MSC_2)\frac{dx_2}{ds} \tag{A3}$$

Equation (A3) says that the welfare change from a small change in the subsidy depends on the quantity response in both markets as well as the relationship between the price and marginal social cost (Oum and Tretheway, 1988; Reguant, 2019). If, for example, the price in the peak market, which represents the marginal willingness to pay in that market, exceeds the marginal social cost in that market, then decreasing consumption by a small amount will decrease welfare (Hahn and Metcalfe, 2021). The same analysis holds for the off-peak market, except the effective price in that market is the subsidized price.

The derivation of equation (A3) makes use of the envelope theorem for the consumer maximization problem. When the off-peak subsidy increases, consumers may alter their travel demand because of an equilibrium change in the subsidy, the price or the congestion level, but the resulting changes in consumption do not a have direct impact on consumer utility. This is consistent with the sufficient statistics approach (Harberger, 1964; Chetty, 2009; Jacobsen et al., 2020).

To evaluate the overall welfare effect of a price subsidy, s', we integrate equation (A3) from $s = 0$ to $s = s'$. If the change in output with respect to a change in the subsidy is constant (i.e., $\frac{dx_1}{ds}$ and $\frac{dx_2}{ds}$ are constant), then the per capita welfare effect of the subsidy is:

$$\frac{1}{N}(W(s') - W(0)) = \int \frac{1}{N}\frac{dW(s)}{ds}ds = (p - MSC_1)\frac{dx_1}{ds}s' + (p - \frac{1}{2}s' - MSC_2)\frac{dx_2}{ds}s' \tag{A4}$$

Equation (A4) (which is the same as equation 1 in the main text) represents the welfare effect of an off-peak subsidy, s', on a rider. It allows for the rider to respond to a change in the equilibrium level of congestion, and the for price to be above or below the marginal private cost. The change in total welfare depends on the difference between price and the marginal social cost, the subsidy, and the changes in demand response. The intuition is similar to that

[51] We consider the implications of relaxing the assumption on off-peak external costs in appendix A.3

developed above for equation (A3).[52] However, the expression for price in the off-peak period, $p-s$, is replaced by the average price as the subsidy varies from 0 to s' (i.e., $(p-\frac{1}{2}s')$).

We implement equation (A4) using causal estimates of demand that compare the case of exogenous congestion with endogenous congestion. This is because in some applications people may take account of their impact on congestion in their actual demand, and in other cases they may not. Thus, it would be useful to know how introducing endogenous congestion affects welfare. In general, the answer is that the impact on welfare in moving from exogenous congestion to endogenous congestion is ambiguous. However, with a few additional assumptions, it is possible to sign the impact of moving from exogenous congestion to endogenous congestion in the peak and off-peak markets separately. Specifically, in our application, we can show that moving to endogenous congestion will generally reduce losses in the peak period and reduce gains in the off-peak period. The overall impact on welfare will depend on whether the reduction in losses in the peak period exceeds or falls short of the reduction in gains in the off-peak period.[53]

Graphical analysis of welfare

The expression for welfare, equation (1), lends itself to a graphical interpretation. Suppose, for simplicity, that demand is linear in the peak and the off-peak market and that p exceeds MSC_1. Figure A1 shows the impact on welfare in the peak period of moving from a subsidy of 0 to a subsidy of s'. The effect of the subsidy is to shift demand in from D_0 to D_1, with a reduction in consumption from x_1 to x_1'. The net welfare loss is given by the negative of rectangle C (which corresponds to $(p-MSC_1)\frac{dx_1}{ds}s'$ in equation (1)). It consists of a decline in producer profits (C+E) and a reduction in congestion costs, E. The resulting change in welfare is –(C+E) + E = -C.[54] That is, welfare declines in the peak market with the introduction of the subsidy in this example.

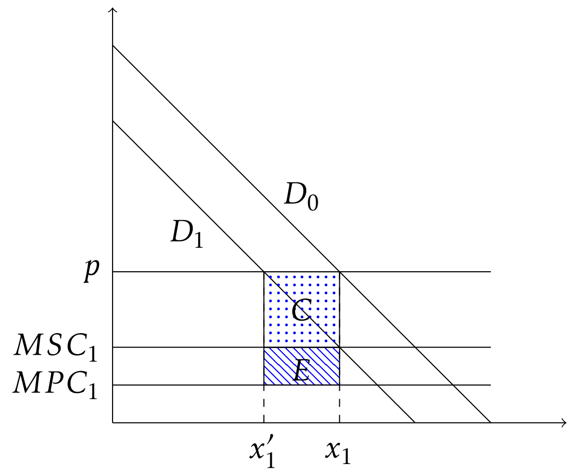
The graph for the off-peak market is shown in Figure A2. The off-peak demand is given by D. The subsidy price is reduced from p to $p-s$, with an increase in demand from x_2 to x_2'. Because there is no change in congestion in the off-peak period (it is presumed to be zero), we

[52]It is of some interest to understand what would happen to welfare if prices in different markets were set "optimally" in the sense that price equals marginal social cost in those markets. A straightforward extension of our model provides an intuitive answer. The introduction of a subsidy leads to a reduction in welfare when prices are set optimally because it distorts consumption patterns away from the optimum, giving rise to a fiscal externality that reduces welfare. To see this result, we need to revise our model slightly. Instead of setting a price, p, in both markets, suppose the regulator could set a price of p_1 in the peak market and p_2 in the off-peak market. Then if p_1 were set equal to MSC_1 and p_2 were set equal to MSC_2, the expression in (A4) for the per capita welfare effect is $-\frac{1}{2}s'^2\frac{dx_2}{ds}$. Assuming that off-peak travel increases with an off-peak subsidy implies welfare declines. That is, an off-peak subsidy adversely effects welfare if prices in the off-peak and peak markets are set to their marginal social cost.

[53]See appendix A.16 for further details.

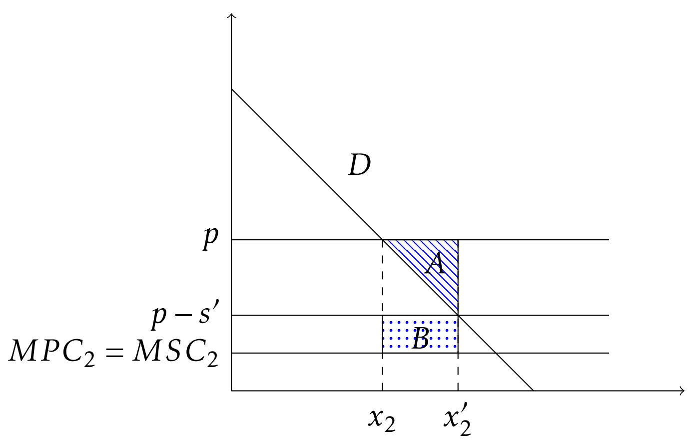
[54]We ignore any direct changes in consumer utility resulting from the subsidy in Figure A1 and Figure A2 below because of the envelope theorem.

Figure A1: Graphical illustration of welfare impact of off-peak subsidy on peak period consumers



Notes: The welfare loss from the subsidy is given by rectangle C in this example. It reflects a combination of the fiscal externality change due to a reduction in producer profits, -(C+E), and a reduction in the congestion externality (E).

Figure A2: Graphical depiction of welfare impact of off-peak subsidy on off-peak period consumers



Notes: The increase in welfare is given by A+B. See text for details.

can focus on the net change in financial flows from the government to compute the change in welfare. The change in profits for producers from the additional units sold $(x_2' - x_2)$ is given by the rectangle $B + 2A$. This means the government can reduce its expenditures to cover BART's losses by that amount. The change in the fiscal externality as the subsidy is gradually increased from 0 to s' is given by triangle A, which increases the government's cost.[55] The welfare change in the off-peak period is the difference between the increase in producer profits (B+2A) less the change in the fiscal externality (A), giving A+B (which corresponds to $(p-\frac{1}{2}s'-MSC_2)\frac{dx_2}{ds}s'$ in equation (1). Adding this off-peak welfare change (A+B) to the change in the peak period (-C) gives A+B-C, which corresponds to equation (1).

Modeling the response to congestion: endogenous and exogenous

An important issue for welfare is how consumers respond to the subsidy. We consider the two polar cases of exogenous and endogenous congestion. These two cases will typically result in different values for $\frac{dx_1}{ds}$ and $\frac{dx_2}{ds}$ in equation (A4), and thus, different empirical measures of the welfare impact of the subsidy. In particular, if a rider takes others' choices to shift to off-peak into account, in equilibrium, fewer riders may shift to off-peak for a given subsidy.

To see how the two cases relate formally, define the peak and off-peak demand response of riders to congestion as $\frac{\partial x_1(p,s,e)}{\partial e} = v_{1e}$ and $\frac{\partial x_2(p,s,e)}{\partial e} = v_{2e}$. In the exogenous congestion case, $v_{1e} = v_{2e} = 0$ (i.e., riders would not adjust their demand in response to a change in the level of congestion (holding constant prices p and subsidy s); In the endogenous case, $v_{1e} \neq 0$ or $v_{2e} \neq 0$. We assume that the price effect of the off-peak subsidy is to increase off-peak demand ($\frac{\partial x_2}{\partial s} > 0$) and reduce peak demand ($\frac{\partial x_1}{\partial s} < 0$). If a higher level of congestion reduces peak demand and increases off-peak demand when price is held constant ($v_{1e} < 0, v_{2e} > 0$), the equilibrium change in peak and off-peak travel in response to an off-peak subsidy ($\frac{dx_1}{ds}$) and ($\frac{dx_2}{ds}$) would be smaller in *magnitude* with endogenous congestion than with exogenous congestion (when $v_{1e} = v_{2e} = 0$) (see appendix A.16 for a formal proof.)

The intuition is that consumers will be less responsive to the price change when they take into account the actual level of congestion on the network. With endogenous congestion, people not only observe the subsidy and react, but also react to the levels of congestion, which have gone down. They therefore do not switch as much in the endogenous case because the effective price of congestion has gone down (compared to the exogenous case).

A.2 Optimal level of the subsidy

We derive the optimal level of the off-peak subsidy for the natural experiment. The optimal level of the off-peak subsidy can be found by setting $\frac{dW(s)}{ds} = 0$, *i.e.*, setting equation A3 equal to zero.

This implies the optimal subsidy, s^*, is given by:

[55] This triangle represents the change in subsidy expenditure resulting from the off-peak consumption change.

$$s^* = (p - MSC_1)\frac{\frac{dx_1}{ds}}{\frac{dx_2}{ds}} + (p - MSC_2) \tag{A5}$$

if $(p - MSC_1)\frac{dx_1}{ds} + (p - MSC_2)\frac{dx_2}{ds} > 0$, otherwise $s^* = 0$.[56]

A.3 Extension of model for welfare analysis of the field experiment

In this section we extend the model in Section 2, allowing for multiple time periods and heterogeneous groups of riders. Each group differs by their route, the route's congestion level, and the time of departure. As discussed, we exploit the fact that the field experiment had random price variation at different times in the morning to identify key parameters.

Assume there are T time periods, J groups of riders, each with population size n_j. The rider population is $N = \sum n_j$. There are $T * J$ different congestion levels in terms of crowdedness. A level of congestion for group j in period t, e_{jt}, is assumed to only affect those riders. Fares for particular groups, denoted by p_j, may vary based on the distance each group travels.

Policy makers use a vector of subsidies, $\mathbf{s}$, which can vary by time period and group. A representative element s_{jt} denotes the subsidy for group j at time t.

Consumers are assumed to be identical within a group, and each consumer maximizes their quasi-linear utility function:

$$u_j\left(x_j, e_j\right) + y_j$$

subject to the budget constraints $\sum_t \left(p_j - s_{jt}\right)x_{jt} + y_j = z_j$, where $\boldsymbol{x_j}$ is a vector of demand for group j on each of the time period. $\boldsymbol{e}$ is a vector of congestion levels for all the time periods and groups. Demand from group j for traveling at time t is $x_{jt}(p, s_j, e)$, which gives the indirect utility function $V_j(p, s_j, e)$. Total demand in the system at time t is $x_t\,(p,s,e) = \sum_j x_{jt}(p, s_j, e)n_j$.

The profit of the firm (which can be negative) equals:

$$\pi = \left(\sum_t p_j x_t\,(p,s,e)\right) - C\left(x_1(p,s,e),\ldots,x_T(p,s,e)\right)$$

where $C(x_1,...,x_T)$ is the cost function for BART. We further assume marginal private cost (per trip) does not differ by j or t, which we refer to as $\frac{dC}{dx_{jt}} = MPC$ for any j and t. Subsidy

[56]This formulation assumes that the envelope theorem holds and the slopes of the demand curves are constant with respect to a change in the subsidy. In principle, one could accommodate changing demands with a more complicated formula.

expenditure equals the sum given to each of the group:

$$\sum_j n_j \left(\sum_t s_{jt} x_{jt} \right)$$

Social welfare is the sum of utility of each groups of riders, plus profit of the firm minus the subsidy expenditure.

The social welfare, $W(s)$, is a function of the subsidy vector $\boldsymbol{s}$

$$W(\boldsymbol{s}) = \sum_j n_j V_j \left(p, s_j,\ e\right) + p\left(\sum_t x_t\right) - C\left(x_1, \ldots, x_T\right) - \sum_j n_j \sum_t s_{jt} x_{jt} \tag{A6}$$

To simplify, we assume congestion in time period t for group j , e_{jt}, affects only the utility of group j, but not other groups. This is a reasonable assumption when a crowded train on a specific route affects passengers travelling on that route at the specific time, but does not affect passengers who travel on other routes or other time.

The welfare effect of a marginal increase of subsidy s_{jt}, for group j to travel in time period t is:

$$\frac{dW}{ds_{jt}} = n_j \left(\left(p_j - MPC\right) \left(\sum_{k=1,..,T} \frac{dx_{jk}}{ds_{jt}} \right) - \sum_{k=1,..,T} s_{jk} \frac{dx_{jk}}{ds_{jt}} \right) + \sum_{k=1,...,T} \phi_{jk} \frac{dx_{jk}}{ds_{jt}}$$

where $\phi_{jk} = \frac{\partial u_j}{\partial e_{jk}} \frac{\partial e_{jk}}{\partial x_{jk}}$ is the marginal external cost of a trip by group j in time period k, analogous to MEC we presented in section 4.

We assume each group receives a subsidy for one period of time. Therefore, we could re-index the time periods for each of the group j, such that the time-period t the group usually travels at as $t = 1$, representing “peak". For the period that they receive a positive subsidy, we refer it to $t = 2$, or interchangeably $t(j)$ to indicate that it depends on each of the group j, representing “off-peak". We further assume that the demand response for group j in other time period are all zero, i.e. $\frac{dx_{jk}}{ds_{jt}} = 0$ for $k \neq 1, 2$, and riders do not change their travel in response to change in congestion level (i.e. “exogenous congestion” case). Following our assumption in section 2, we assume the demand response is locally constant (a linear demand curve).

The effect of subsidy to group j, analogous to equation (A3), is therefore

$$\frac{dW(s)}{ds_{jt(j)}} = n_j \left(\left(p_j - MPC + \phi_{jt}\right) \left(\frac{dx_{j1}}{ds_{jt}} + \frac{dx_{jt}}{ds_{jt}} \right) - s_{jt} \frac{dx_{jt}}{ds_{jt}} + \left(\phi_{j1} - \phi_{jt}\right) \frac{dx_{j1}}{ds_{jt}} \right)$$

Denote the total welfare change is $\Delta W_{jt}\left(s_{jt}\right)$ for a subsidy of size s_{jt} , the welfare effect per participants in group j is,

$$\frac{1}{n_j}\Delta W_{jt}\left(s_{jt}\right) = \frac{1}{n_j}\int_0^1 \frac{dW}{ds_{jt(j)}}\frac{ds_{jt(j)}}{d\theta}d\theta$$
$$= -\frac{1}{2}s_{jt}^2\frac{dx_{jt}}{ds_{jt}} + \left(\phi_{j1} - \phi_{jt}\right)\frac{dx_{j1}}{ds_{jt}}s_{jt} + \left(p_j - mc + \phi_{jt}\right)\left(\frac{dx_{j1}}{ds_{jt}} + \frac{dx_{jt}}{ds_{jt}}\right)s_{jt} \quad \text{(A7)}$$

We define MSC_1 and MSC_2 (marginal social cost of the two periods) by $MSC_1 = MPC + \phi_{j1}$, $MSC_2 = MPC + \phi_{jt}$, similar to section 2 and with the same interpretation. Equation A7 could also be rearranged in terms of MSC_1 and MSC_2,

$$\frac{1}{n_j}\Delta W_{jt}\left(s_{jt}\right) = -\frac{1}{2}s_{jt}^2\frac{dx_{jt}}{ds_{jt}} + (p_j - MSC_1)\frac{dx_{j1}}{ds_{jt}}s_{jt} + (p_j - MSC_2)\frac{dx_{jt}}{ds_{jt}}s_{jt} \quad \text{(A8)}$$

which is the expression presented in the theory developed in section 5. We use equation A8 to evaluate the welfare effect of each of the subsidies provided in the field experiment. It gives the welfare effect per person who received the subsidy for a particular group.

A.4 Measuring the marginal external cost of a trip for the natural experiment

To estimate the marginal social cost of a trip in the peak period (MSC_1), we need an estimate of marginal external cost (MEC_1). We obtain MEC_1 by making assumptions about the congestion technology, the value people place on congestion in the peak period, and their value of time (VOT).

We define congestion, e, as crowding (or density) and delay. Crowding is measured as the number of passengers per square meter in a train car in the peak period. The reason we chose this measure is because reducing crowding was an objective in both experiments and it is an important externality. We use train level crowding data from BART to calculate the average density of a train car.

The marginal external cost during the peak period is defined as $MEC_1 = \frac{\partial u}{\partial e}\frac{\partial e}{\partial x_1}$. It is the product of the effect of a congestion change on utility and the effect of an increase in peak rides on congestion (crowding measured in density).

We assume an additional peak period trip has a constant effect on congestion that does not vary with the level of subsidy. Moreover, the value placed on incremental congestion in the peak period by the representative rider is also assumed to be constant. This implies MEC_1 is also constant.

We can derive the marginal external cost in the peak period with some additional assumptions. We assume passengers are distributed evenly across all train cars in the system, therefore $n = \frac{Nx_1}{k}$ where k is number of train cars during peak period, n is the number of passengers per car, and $e = \frac{n}{a} = \frac{Nx_1}{ka}$ is the density in the car with a as the size of a car.[57] Under this assumption, the effect of a peak period trip on density in a train car is $\frac{\partial e}{\partial x_1} = \frac{e}{x_1}$.

We define the effect of a congestion change on utility that results from crowding and delay in journey time as ϕ_c and ϕ_d respectively.

A rider's disutility from crowding depends on the level of congestion. We assume a minute on a train with congestion level e gives rider disutility equivalent to that of spending $T(e)$ minutes on a train with $e = 0$. (i.e., $T(e)$ is the "time-multiplier" in Haywood and Koning (2015)). Therefore, the disutility from crowding for a train ride of t minutes with congestion level e equals $t * T(e)$ in terms of time, and $VOT * t * T(e)$ in monetary terms, where VOT is the value of time for a rider. If riders on average take x_1 trips a day on a congested train (e.g., in the peak period), the disutility from crowding per day is $VOT * t * T(e) * x_1$.

We assume that a marginal increase in density e increases the (per minute) disutility from

[57] We relax this assumption in our field experiment. See discussion below.

crowding by a constant factor k (i.e. $T^{'}(e) = k > 0$).The change in utility resulting from a marginal increase in congestion therefore equals $\phi_c = VOT * t * k * x_1$, holding VOT, t and x_1 constant.[58] Similarly, a rider may be delayed when congestion increases because it takes longer for all passengers to board the train. We assume an increase in density increases journey time by a constant, o. We can empirically estimate o from our data (see below). For each trip, this results in a monetary loss of $VOT * o$. Since each rider takes x_1 peaks trip per day, $\phi_d = VOT * o * x_1$.

Following Haywood and Koning (2015), and adding the effect from delay, we define

$$MEC_1 = \frac{\partial u}{\partial e}\frac{\partial e}{\partial x_1} = (\phi_c + \phi_d)\frac{\partial e}{\partial x_1} = VOT * t * k * e + VOT * o * e \tag{A9}$$

where the second equality is a first order approximation.

We assume the value of time is half the hourly wage (Small, 2012) or 75% of the hourly wage (Goldszmidt et al., 2020), which implies VOT = \$12.5 or VOT = \$18.68 respectively (the median wage in San Francisco in 2016 was \$24.9).[59] We calculate the average in-train time in the peak period as 37 minutes from our BART gate entry and exit data.

The average density in the peak period, e, is 1.13 passengers per square meter, estimated from our novel BART train crowding data. We use a time-multiplier of 0.11 from Haywood and Koning (2015). Thus, $VOT * t * k * e = \$0.954$.

We find that an extra passenger per car results in a delay in departure time relative to its schedule time of 1.37 seconds (see Table A1), or 0.023 minute (=1.37/60). This implies a unit increase in density increases journey time by 1.59 minutes on average, (as $d = \frac{n}{a}$, where d is density, n is number of person per car, a is the size of the car (68.93 sqm), let t be the journey time in second, $\frac{\partial t}{\partial d} = a\frac{\partial t}{\partial n}$ = 0.023*68.93) i.e., o = 0.026 (= 1.59/60mins). Thus, $VOT * o * e$ = \$0.37 (= \$12.5*0.026*1.13). This implies MEC_1 = \$ 1.33 per peak hour trip.

To estimate the impact of an extra passenger per car on the delay, that allows us to calculate the effect of an increase in density on delay (i.e., o), we use train-level data from December 2016 to March 2017. The data contains the scheduled and actual departure information

[58] As our model in section 2 is in terms of representative rider, the fact that ϕ_c, the change in utility resulting from a marginal increase in congestion depends on the number of peak ride per person x_1, is equivalent to that the marginal change in congestion density is increasing with the total number of rides in the system. We assume that the marginal external cost of a ride in off-peak to be zero in section 2, this could be considered as either because T'(e) to be very low at a low level of e, i.e. a change of density has no impact on the utility of a rider per ride because of preferences. Otherwise, the marginal external cost in the off-peak period could be lower than the peak-period because there are fewer riders experience following equation A9, and it would only be approximate zero if the density/traffice in the off-peak period is close to zero.

[59] Metropolitan and Non-metropolitan Area Occupational Employment and Wage Estimates, San Francisco-Oakland-Hayward, CA, US Bureau of Labor Statistics, May 2016

for each train at each station, for all trains that run for the day. We estimate the following specification using the train and station level data

$$d_{ijkdt} = \alpha + \beta n_{ijkdt} + \mu_{ij} + \gamma_t + \eta_d + \epsilon_{ikdt} \tag{A10}$$

where d_{ijkdt} is the difference between the actual departure (door closing) time with the scheduled departure (door closing) time, for train k on date d scheduled at time t departing from station i towards station j. n_{ijkdt} is the number of passenger per car on the train. To control for any systematic delay at the station or route level, we control for the station-direction specific fixed effect μ_{ij}. We control for 20 minute time interval fixed effects γ_t, and date specific fixed effects η_d that account for unobserved factors that affect delay on specific dates or time. As robustness check, we control for any planned delay in terms of scheduled run time between the previous station and station i for the train k (X_{ijkdt}).

The number of passenger per car for a train at a station i is related to the net number of passengers getting on the train (i.e., number of people boarding the train net number of people getting off the train) at all the stations before i, plus that at station i. To address potential confounding factors at the route-time level, e.g., unexpected events or accidents happening at the route-level or at the stations before i that affects both the delay time and the density, we use the number of passenger boarding the train cars at station i as instrument for the passenger per car on the train at station i (before departure). Our estimate therefore exploit the variation in the number of boarding passengers at different stations and time to identify the effect of congestion on delay. The OLS and IV estimates of β in the above equation are in Table A1.

An increase in density in a train car, related to an extra passenger boarding at a specific station, could be associated with additional delay because of: i) that resulting from when the passenger board at the origin, ii) an increase in in-car crowdedness slowing down boarding of other passengers at stations in between the origin and destination, and c) that resulting from when the passenger leave the train at the destination. We estimate equation A10 exploiting empirically the variation in number of passenger boarding each station, it is therefore most closely represent the effect of a passenger boarding the train on delay (i.e. the effect of the density on delay estimated in equation A10 is local with respect to the instrument). We do not assume that the passenger would add to further train delay in stations beyond the origin, e.g. if there is no further passenger boarding at later stations, it is unlikely that a passenger in-car would add to further delay incrementally at each stations.

Table A1: The effect of density on train delay

	Outcome: Train-station delay (sec) Sample - Peak period trains			
	(1)	(2)	(3)	(4)
Passenger per car	0.636***	0.504***	1.370***	1.388***
	(0.153)	(0.180)	(0.344)	(0.343)
Observations	72181	71540	72181	71540
Train-station specific controls	N	Y	N	Y
Date FE	Y	Y	Y	Y
Time-specific train FE	Y	Y	Y	Y
Station-line FE	Y	Y	Y	Y

Notes: The table presents the regression results with the outcome as the difference between a schedule departure time (door closing) and the actual departure time (door closing) on the passenger per train car on the train. Controls includes date fixed effect.

A.5 Estimation of the marginal cost per trip

We estimate the marginal private cost of a BART trip using information on BART's operating costs. We combine various sources of BART financial and operations information to compute the marginal private cost of an individual BART trip. Our measure of operating expenses includes the cost of maintenance.[60]

Assuming the total operating cost of BART depends only on the total number of passenger trips, i.e., $C(Nx_1, Nx_2) = C(Nx_1 + Nx_2) = C(Nx)$, the relationship of marginal cost to average cost with respect to the number of passenger trips can be estimated using the following equation:

$$C(Nx)_t = \alpha + \beta Nx_t + \epsilon_t \tag{A11}$$

where $C(Nx)_t$ is the total operating cost of BART in year t, Nx_t is the annual number of total trip in the system, and β represent the marginal private cost per BART trip.

From 2007 to 2016, the cost per passenger increases from \$4.47 to \$4.91 in nominal terms.[61] In real terms, it declines from \$ 5.66 to \$ 4.91,[62] while the annual ridership increases from 101 to 128 million.[63] Figure 1 suggests that there is closer relationship between annual ridership and total operating cost.

Table A2 column (1) provide estimate of the above equation. The implied marginal cost in

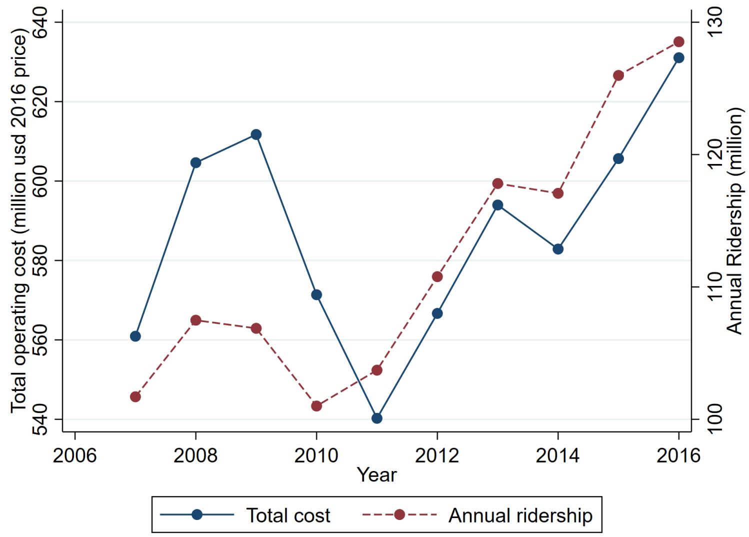
[60] p. 4-20, BART SRTP-CIP.

[61] Figure 3-2 BART SRTP-CIP.

[62] Deflated by San Francisco CPI provided by Bureau of Labor Statistics (https://www.bls.gov/regions/west/data/consumerpriceindex_sanfrancisco_table.pdf).

[63] Figure 3-3 BART SRTP-CIP.

Figure A3: Cost per trip and annual number of trips



Note: The graph plots the total operating cost and annual ridership from 2007 to 2016, in 2016 prices.

2016 is $1.89. The estimate is robust to controlling for yearly average energy price per kwh in San Francisco.

Table A2: Estimation of marginal cost

	(1) totalcost	(2) totalcost
Annual ridership	1.898**	1.695
	(0.724)	(1.184)
Energy price (1000 kwh)		-160996.3
		(712507.1)
Observations	10	10

Note: The table represent estimate of equation A11. Total operating and annual rider ship data from from BART annual financial report. Energy price is annual average price per kwh (in 2016 price), data source from U.S. Bureau of Labor Statistics, Electricity Per KWH in San Francisco-Oakland-Hayward, CA (CBSA) [APUS49B72610], retrieved from FRED, Federal Reserve Bank of St. Louis; "https://fred.stlouisfed.org/series/APUS49B72610", October 21, 2020.

A.6 Estimation of feedback effect from congestion

In this section, we present the methodology and the empirical estimates on the demand response to congestion. When the level of congestion increases or decreases, this may affect the travel demand by consumer, holding other factors constant. We need to estimate this demand parameter to empirically evaluate and scale-up the welfare consequences of small-scale price subsidies when considering the endogenous congestion case.

In the endogenous case, we need to estimate two important parameters, $v_{11} = -\frac{\partial x_1}{\partial e}\frac{\partial e}{\partial x_1}$ and $v_{21} = -\frac{\partial x_2}{\partial e}\frac{\partial e}{\partial x_1}$, which are the demand responses of peak and off-peak travel with respect to a marginal change in peak-hour traffic in the aggregate. Intuitively, it measures how individual demand responds to (expected) aggregate congestion changes. In this section we try to estimate the demand response with respect to the aggregate travel, via congestion as the channel.

We use the travel data of participants in the natural experiment to estimate these parameters. We exploit variation in naturally occurring congestion changes, and that some riders may be more affected by congestion at certain stations, to identify how riders responds to congestion. Our strategy focus on variation in naturally occurring congestion at some part of the system, e.g., idiosyncratic events or conferences near the stations that lead to an increase in demand to travel to the area (therefore increasing congestion), and control for potential confounding factors that affect the demand on all parts of the system, e.g., weather.

The majority of morning peak hour journey in BART uses two exit stations, Embarcadero and Montgomery Street. Riders who regularly uses these busy stations are more likely to be affected by congestion at these stations, and may adjust their travel demand accordingly when (expected) congestion at these stations changes. We estimate how congestion at these two exit stations in the morning affects the peak and off-peak demand for riders who regularly exit at these two stations. We use other riders who do not regularly use these two stations as the control group (in the period before the price subsidies were introduced).

The treatment variable is the aggregate number of rides that exit at these two station on a given day (*i.e.*, comparing naturally occurring busy and quiet days), interacted with a dummy variable in which the rider regularly uses these two stations (comparing with riders who do not regularly uses these two stations). This is effectively a difference-in-differences strategy.

The aggregate number of rides that exit at the two stations proxy for the (expected) level of congestion at the two stations on each day. We assume that riders who travel regularly on route that exit at these two stations have correct expectation on the congestion at these two stations. For example, a rider who travels regularly would be able to infer from public information the level of congestion at the stations they uses regularly. Under this assumption, our strategy estimate the demand respond to a change in the congestion level that are relevant to the riders' trip.

To control for unobserved factors that affects both demand and overall usage of BART system, e.g., weather or events that disrupt or affect travel demand for the whole BART system, we use riders who do not regularly exit at the two stations as the control group. We assume that the demand from riders who do not travel regularly at these stations would not depend on the congestion level on the station that they do not travel with.

Specifically, we estimate the following equation

$$y_{ijt} = \beta Busy_i * PeakBusy_t + \mu_t + \gamma_i + \epsilon_t \tag{A12}$$

where y_{it} is number of peak hour trip by user j on date t. $Busy_i$ is an indicator for user i regularly exit at Montgomery street and Embarcadero station in morning, the two busiest stations in BART. $PeakBusy_t$ is the total number of trips (started) in peak hour that exited at the two busy stations, in the whole BART system. γ_i is rider fixed effects and μ_t is date fixed effects.

Table A3 reports the estimate of equation (A12).The sample include the travel demand of the Perks program participants from April 2016 to February 2017. There are 17,185 riders in the sample, among them, 7,477 use the Embarcadero and Montegomery Street stations regularly–this is defined as these users exiting at one of the two stations with above median percentage (median is around 97%) among their morning trips (5.30-10.30am) in the sample period. We find that an increase in 10,000 peak hour riders exiting at the two busiest station reduces the

Table A3: Peak hour trip by riders who regularly exit at Montgomery street and Embarcadero, compared with others

	Number of trip			
	(1)	(2)	(3)	(4)
	Peak	Peak	Off-peak	Off-peak
Rider who regularly uses Embarcadero and Mont. St	-0.00617*	-0.00635*	0.00858**	0.00877**
× Peak hour rides exiting at E&M (10,000)	(0.00371)	(0.00372)	(0.00365)	(0.00366)
N	1807471	1807471	1807471	1807471
Rider FE	Y	Y	Y	Y
Date FE	N	Y	N	Y

Notes: The table present the estimates of equation A12. Sample includes dates that riders have at least 1 trip for the estimation of intensive margin response. Standard errors robust to 7 days of auto-correlation.

peak demand of riders who uses these routes regularly (among the participants of experiment 1) by -0.00635 per trip per day at peak hour. We also find that their demand for off-peak trip increase by 0.00858 trip per day at off-peak hour.

The above are estimates of the effect of changes in aggregate rides on trip demand per person - $\frac{\partial x_1}{\partial e}\frac{\partial e}{\partial N x_1}$ and $\frac{\partial x_2}{\partial e}\frac{\partial e}{\partial N x_1}$. For the welfare analysis in section 4, we scaled the aggregate rides by the total number of BART riders, which gives the feedback effect of the increase in peak rides per person on the peak and off-peak demand $\frac{\partial x_1}{\partial e}\frac{\partial e}{\partial x_1}$ and $\frac{\partial x_2}{\partial e}\frac{\partial e}{\partial x_1}$. Using $N = 271,341$, we find $v_{11} = -0.172$ and $v_{21} = 0.237$.We use the estimates of these two parameters for our welfare analysis in section 5.

A.7 Additional description for parameters for welfare calculation of experiment 1

The following table A4 summarizes the key parameters used in the welfare calculation in the natural experiment.

Table A4: Additional description for parameters for welfare calculation 1

Parameter and description	Value	Source
s' ($ / trip)	$0.25	Average subsidy per off-peak trip in the Perks experiment.
$\frac{dx_2^*}{ds}$ (change in quantity/ $)	0.0664 (exogenous) 0.0625 (endogenous)	1st Perks experiment ($\frac{dx_2}{ds} = -\epsilon_{22}\frac{x_2}{p}$, $\epsilon_{22} = -0.86$; see Section 3)
$\frac{dx_1^*}{ds}$ (change in quantity/ $)	-0.0194 (exogenous) -0.0166 (endogenous)	1st Perks experiment ($\frac{dx_1}{ds} = -\epsilon_{12}\frac{x_1}{p}$, $\epsilon_{12} = 0.44$; see Section 3)
v_{11} (feedback effect from congestion on peak-ride)	-.172	Detail of estimation in appendix A2
v_{21} (feedback effect from congestion on off-peak ride)	0.237	Detail of estimation in appendix A2
MEC_1 (the external cost is measured in $/peak-trip)	-0.537	See appendix A1 for formula.
t (In-vehicle time)	37 (minutes)	Calculated form BART journey data (gate entry to gate exit)
$Tm'(d)$ (Time multiplier)	0.11 (minutes/passengers per sq m)	Haywood and Koning (2015)
Value of time (VOT)	= 0.116 per minute (= $7 per hour)	Small (2012) - value of time as half of wage. minimum wage in SF in 2017 = $14
d (density in car, number of passengers per sq m)	1.13 (= 77.9/68.93)	Weighted average number of passengers per car/car size (weighted by number of passengers in car)
	n (average passengers per car in peak hour) 41.389	Calculated form BART crowding data
	a (Car size, sq m)	We take type B car as representative car, where the length is 70', with width 10.6'
68.93 (= 70 (feet)*10.6(feet)* 0.3048^2)		
$\frac{de}{ds}$ (change in crowding)	-0.24 (exogenous) -020 (endogenous)	1st Perks experiment.
p (average fare per journey)	$3.99	Average off-peak fare per journey paid by riders in the 1st experiment.
$\frac{dC}{dNx_1}$, $\frac{dC}{dNx_2}$ ($/journey)	$1.89	Calculated using cost and ridership information from BART. See Appendix A5.
N	271,341(people)	Average number of morning peak trip in BART per day (47,756) / peak-trip per rider per days (estimated from data of participants in the first Perks experiment) (0.176)

Table A5: Additional description for parameters for MVPF calculation of experiment 1

Description	Value	Source
Inframarginal benefit (x_2)	0.307	1st experiment
Congestion benefit ($-\frac{dx_1}{ds} * MEC_1$)	0.0258	1st experiment; see Appendix A.3
Change in direct subsidy cost (x_2)	0.307	1st experiment
Change in transfer to BART ($-(p - MPC)(\frac{dx_1}{ds} + \frac{dx_2}{ds})$)	-0.0987	1st experiment estimates
Marginal private cost (MPC)	1.89	See Table A4
Price per trip (p)	$3.99	See Table A4
Marginal external cost (peak period) (MEC_1)	1.33	See Table A4

A.8 Welfare effect of the natural experiment: Transbay and other routes

This section presents the welfare calculations for a targeted subsidy of \$0.25 for the Transbay route (i.e., a heavily congested route) and all other routes. The key point is that welfare increases by about 23% percent when targeting the Transbay route compared with all other routes. We also provide a novel welfare framework that allows for everyone on BART to receive a price subsidy (see Appendix A17).

The welfare calculation assumes that all network effects of the subsidy are captured by our demand parameters derived in section (3), and the two systems can be treated independently. In terms of the welfare equation (equation (1)), we are measuring differences in demand and marginal social cost across the two sets of routes.

The first three rows of the table provide key parameters. MEC_1 and MSC_1 represent the marginal external cost and marginal social cost in the peak period. Congestion is a key determinant of MEC_1 as explained in section A.4. Congestion is measured as the number of passengers per square meter in a train car (see Appendix Table A4 for details). The equilibrium level of congestion depends on whether consumers take congestion into account. The congestion during the peak period is about 0.4% percent higher in the case of endogenous congestion. The welfare benefit associated with congestion alleviation is smaller in the endogenous case because of a range of factors, including a weaker demand response.

The third and fourth rows of the table summarize welfare changes associated with peak and off peak travel. The total welfare change (representing the sum of the welfare changes for the peak and off-peak travel) is given in the final row for various cases.

Table A6: Welfare effect of the subsidy for the Transbay route and other routes

	Transbay routes			Other routes		
Parameter	Exogenous congestion	Endogenous congestion	Percent Change	Exogenous congestion	Endogenous congestion	Percent Change
MEC_1	\$2.23	\$2.23	.	\$0.93	\$0.93	.
MSC_1	\$4.12	\$4.12	.	\$2.82	\$2.82	.
Congestion	1.518	1.5243	0.4	0.8116	0.815	0.4
Welfare change with peak travel change	0.0006	0.0005	-14.7	-0.0057	-0.0048	-14.7
Welfare change with off-peak travel change	0.0328	0.0309	-5.9	0.0328	0.0309	-5.9
Total Welfare change	0.0334	0.0314	-6.1	0.0271	0.026	-4.1

A.9 Optimal subsidy and welfare impact in the natural experiment

Table A7: Welfare and congestion impacts for the optimal subsidy

Parameter	Exogenous cong.	Endogenous cong.	Percent Change
Welfare change associated with peak travel change	-0.028	-0.0239	-14.7
Welfare change associated with off-peak travel change	0.1448	0.1362	-5.9
Total Welfare change	0.1168	0.1123	-3.8
congestion	0.6621	0.6874	3.8
Optimal subsidy	1.875	1.896	1.1

Notes: The table calculates the optimal subsidy using equation A5 in section A.2, for the natural experiment. It uses parameters described in section A.7 for the baseline case. The optimal subsidy is over seven times higher than the actual subsidy in the experiment. Overall welfare increases with the optimal subsidy. The decline in ridership in the peak period reduces welfare, and the increase in ridership in the off-peak period increases welfare. See text for details.

A.10 Additional discussion for welfare analysis

This section explains how we include the benefits from reduced externalities from other modes of transit in an illustrative calculation.

Including externalities in other modes of transport

This section analyzes the potential impact of other travel modes on the estimation of welfare from section 5. The basic point is that considering possible travel shifts from other more polluting modes to mass transit could increase the welfare estimate. We estimate this increase in welfare could be 15.4 percent.

While our model in section 2 does not explicitly consider other modes of transport, it is possible that other modes of transport, such as automobiles or bus, could affect welfare. Using the sufficient statistics framework, if the price paid by the consumer is equal to the marginal social cost in the other mode of transport, then there would be no change in welfare resulting from the consumption change in that mode.

If these other modes of transport are not priced optimally, the change in consumption in the other modes would have an impact on the welfare. Our estimate of the welfare effect of the subsidy could be biased downward if the market prices in these other modes are below their marginal social cost. This is because we find net increases in travel on BART with the introduction of the subsidy. Formally, we consider the case where $\frac{dx_1}{ds} + \frac{dx_2}{ds} > 0$, which implies, on average, that there are participants who demand more BART rides after receiving the subsidy. This is consistent with our empirical estimation in the natural experiment.

We provide an illustrative calculation for the case of emission reductions that could result from the mode shift to BART.[64]

For simplicity, we will assume that total travel remains unchanged and an increase in a BART trip reduces automobile travel by the same amount in terms of miles traveled. Then, the net increase in welfare from shifting travel is given by the following formula: $\left(\frac{dx_1}{ds} + \frac{dx_2}{ds}\right)s'\phi \geq 0$, where ϕ_r is the externality measured in dollars per trip.

Using the parameters in Table A8, we calculate that the externality associated vehicle travel per mile associated with air pollution is \$0.0055 per mile of automobile travel. With an average BART journey of 14.6 miles per trip, this implies a net increase in BART trips is associated with a reduction of external cost of \$0.08, *i.e.*, $\phi = 0.08$. We present the welfare estimates with and without the benefits from emission reductions for the natural experiment

[64] It is possible that there could be congestion benefits as well for automobile travelers. We do not include such benefits because we do not have good data on their magnitude for shifts in particular time periods. If, for example, shifts in automobile travel to BART occurred during periods in which there was little congestion on highways, the benefits from reduced congestion would be small.

in Table A9. The table shows that welfare benefits increase by about 15.8 percent in the endogenous congestion case and 15.4 percent in the exogenous case when mode shifting is accounted for.

Table A8: Parameters for vehicle externality

	CO	HC	NO_x	CO_2
Tier 2 Federal Exhaust Standards (2007-2016), grams per mile	3.40	0.10	0.14	
	0.50	0.50	0.50	
Emission per mile (grams)	1.70	0.05	0.07	404
Damage per ton ($)	1045	15047	35566	46.27
Damage per gram ($)	0.0012	0.0166	0.0392	0.000051
Damage per mile ($)	0.0020	0.0008	0.0027	0.0206
Total damage per mile ($)	0.0261			
Average trip length (miles)	14.60			
Average damage per trip	0.38			

Note: This table shows the 'average damage per trip' parameter we have used to calculate the externality that is associated with each extensive margin response per trip from vehicle. It assumes "New vehicle emissions are typically 40 to 50 percent of exhaust standards" following Jacobsen et al. (2022). It does not consider variation in vehicle ages and types.

Table A9: Welfare impacts of emission benefits for the base case and a targeted subsidy

	Parameter	Base case		Targeted subsidy	
		Exogenous congestion	Endogenous congestion	Exogenous congestion	Endogenous congestion
(1)	Total Welfare change	0.0291	0.0277	0.0334	0.0314
(2)	Total Welfare change including vehicle externality	0.0336	0.0321	0.0379	0.0358
	Percent change (%)	15.4	15.8	13.5	14

Notes: Welfare is measured in dollars per person per weekday. Overall welfare increases with the subsidy. The decline in ridership in the peak period reduces welfare, and the increase in ridership in the off-peak period increases welfare. See text for details. The total welfare change is calculated from equation (1) in the first row; and it is calculated with an adjustment factor of $\left(\frac{dx_1}{ds} + \frac{dx_2}{ds}\right)s'\phi$ assuming and $\phi = 0.38$, $s' = 0.25$.

A.11 Additional estimates on welfare impact

In this section we provide a numerical example for how we calculate the annual welfare benefits from the subsidy if all BART riders receive the price subsidy used in the natural experiment.

Calculation of welfare impact

We provide details on how we calculate the annual welfare impact of the subsidy for two cases that help bound the welfare estimates. First we compute the total benefits from the experiment in the case of exogenous congestion if all riders in BART received the subsidy and had the same elastcities as those in the natural experiment.[65] Second we compute the total benefits from the experiment in the case of exogenous congestion if only those riders in BART receive the subsidy, and other riders do not change their behavior.

To compute annual benefits from rolling out the subsidy to all riders, we multiply three numbers – the number of riders, the average daily benefits for those riders, and the number of days in a year. Our estimate of total riders in BART is $N = 271,341$. The average welfare impact per person per weekday is \$0.029 in the base case. This gives an annual welfare estimate of $2,052,966$ (i.e., $0.029 * 271,341 * 260$).

We next compute annual benefits from rolling out the subsidy to a subset of riders who actually participate in the experiment. We use equation (A15), which considers two types of riders. In addition, we make the assumption that riders not in the experiment do not change their behavior. We do this to illustrate the idea in the extreme case, and thus provide a plausible lower bound for the welfare result. In terms of equation (A15), this means that $k = 0$ (i.e., those riders not in the experiment do not change their behavior.) Applying the equation, we let $n_a = 17,545$, $n_b = N - n_a = 253,796$ and other parameters are the same as in the previous case. This gives an average welfare impact of \$0.00188 per person per day. The aggregate welfare impact for the program rolled out for 6 months (130 days) is calculated as $0.00188 * 271,341 * 130 = \$66,373$. If the subsidy program is rolled out for the whole year, the overall welfare impact would be twice that or 0.18 million.

Calculation of net welfare change per dollar of subsidy

In section 4 we divided the net welfare change by the subsidy expenditure to obtain the net welfare change per dollar of subsidy measure. We provide the calculation here.

Under the assumption of exogenous congestion, the net welfare increase for the natural experiment is 0.0291\$ per person per day. We calculate the subsidy expenditure in this case

[65] The calculation of the endogenous congestion case is similar, except we use the numbers for demand responsiveness associated with that case.

as $x_2' * s$, where x_2' is the off-peak travel per day per person with the subsidy, and $s = 0.25$. In this case, $x_2' = x_2 + \frac{dx_2}{ds}s = 0.3236$. This implies that the net welfare benefit per dollar of subsidy expenditure is $\frac{0.0291}{0.3236*0.25} = 0.3597$. If we allow for the two types of consumers to differ in their demand response, and assume that other BART riders are half as responsive to price as the participants in experiment 1, the net welfare benefit is 0.0155 per person per day. $x_2' = x_2 + (\frac{n_a}{n_a+n_b} + \frac{n_b}{n_a+n_b}0.5) * \frac{dx_2^b}{ds}s = 0.316$, where n_a is the number of participants in experiment 1 and n_b is the number of other BART riders. Thus, the net welfare benefit per dollar of subsidy is 0.196.

A.12 Natural experiment welfare estimates when the marginal social cost varies

A key assumption in the body of the paper is that the marginal social cost is constant in the peak period and in the off-peak period. In this appendix, we consider the empirical implications of relaxing that assumption for the peak period. We compare the welfare result using different functional forms for the MEC function. Our main conclusion is that in this particular example, the results do no change much in moving from a constant MEC function to a linear MEC function, or in moving from a constant MEC function to a quadratic MEC function.

In the first subsection, we extend the theoretical analysis contained in Section 2. We start with the general case, which includes the non-linear case, the linear case and the constant case. We then examine the specific case where we impose the linearity assumption, which results in a slightly simpler formula. In the second subsection, we discuss the empirical welfare implications.

A.12.1 Allowing the marginal social cost to vary with the level of congestion

Following the model in Section 2 and equation (1), we derive the formula in the case when the MSC is a non-linear function of congestion, and discuss how that allows us to estimate the impact of MSC_1 varying with e. We present the derivation here. Equation (1) is reprinted below.

$$\begin{aligned}\frac{1}{N}\left(W\left(s'\right)-W(0)\right) &= \int \frac{1}{N}\frac{dW(s)}{ds}ds = \int \left(p-MSC_1\right)\frac{dx_1}{ds}+\left(p-s'-MSC_2\right)\frac{dx_2}{ds}ds \\ &= \int p\frac{dx_1}{ds}s' - \int MSC_1\frac{dx_1}{ds} + \int ds + \left(p-s'-MSC_2\right)\frac{dx_2}{ds}ds\end{aligned} \tag{A13}$$

where the limits of integration are 0 and s'.

When MEC_1 varies with the level of e, we have $MSC_1(e) = MPC + MEC_1(e)$. We can rewrite the second term in the final expression in (A13) as

$$\int_0^{s'} MSC_1(e)\frac{dx_1}{ds}ds = \int_0^{s'}\left(MPC+MEC_1(e)\right)\frac{dx_1}{ds}ds = MPC\frac{dx_1}{ds}s' + \frac{dx_1}{ds}\int_0^{s'} MEC_1(e)ds$$

One way to simplify the preceding expression is to define an average MEC_1 over the range of the subsidy as $\overline{MEC}_1 = \int_0^{s'} MEC_1(e)ds\frac{1}{s}$. The above equation can then be rewritten as:

$$MPC\frac{dx_1}{ds}s' + \frac{dx_1}{ds}s'\int_0^{s'} MEC_1(e)ds\frac{1}{s'} = MPC\frac{dx_1}{ds}s' + \frac{dx_1}{ds}s'\overline{MEC}_1 \tag{A14}$$

(A14) is general, and thus allows for a non-linear functional form for MEC_1. The difference between the second term in equation (A13) associated with the marginal social cost, compared with that of equation (1), is that the constant MEC_1 is replaced with $\overline{MEC}_1$ for the calculation of MSC_1.

The other terms in equation (A13) would be the same as in equation (1) in the text. Estimating $\overline{MEC}_1$ requires knowledge of both the shape of the $MEC_1(e)$ function and also how congestion varies with the subsidy. In particular,

$$\overline{MEC}_1 = \int_0^{s'} MEC_1(e)ds\frac{1}{s'} = \int_{e(0)}^{e'} MEC_1(e)\frac{1}{\frac{de}{ds}}de\frac{1}{s'} \tag{A15}$$

where e' is the congestion at the subsidy level s', $e(0)$ is the congestion level before the subsidy. For example, if we assume $MEC_1(e) = \beta_0 + \beta_1 e + \beta_2 e^2$ and $\frac{de}{ds}$ is constant, we could calculate $\overline{MEC}_1$ numerically.

We can derive a simpler formula if we assume MEC_1 is linear. Suppose we assume that the marginal external cost can be written as $MEC_1(e) = \alpha_0 + \alpha e$, where α_0 is a non-negative constant and α is a positive constant. The second term in equation (A13), associated with the marginal social cost in the peak period, becomes

$$\begin{aligned}\int MSC_1(e)\frac{dx_1}{ds}ds &= \int (MPC + \alpha_0 + \alpha e)\frac{dx_1}{ds}ds \\ &= MPC\frac{dx_1}{ds}s' + \frac{dx_1}{ds}\alpha_0 s' + \frac{dx_1}{ds}\alpha\int_0^{s'} e(s)ds\end{aligned} \tag{A16}$$

If we assume $\frac{de}{ds} = \frac{de}{dx_1}\frac{dx_1}{ds}$ is approximately constant, i.e. $\frac{de(s)}{ds} \approx \frac{e(s')-e(0)}{s'}$, we can simplify (A16) by noting

$$\int e(s)ds = [e(s)s]_{s=0}^{s=s'} - \int sde(s) = e(s')s' - \frac{de(s)}{ds}\int sds = e(s')s' - \frac{(e' - e(0))}{s'}\frac{1}{2}s'^2$$
$$= \frac{1}{2}(e(s') + e(0))s' = \bar{e}s'$$

where $\bar{e}$ is the average of the congestion levels before and after the subsidy change. This implies

$$\int MSC_1(e)\frac{dx_1}{ds}ds = (MPC + \alpha_0 + \alpha\bar{e})\frac{dx_1}{ds}s' = MSC_1(\bar{e})\frac{dx_1}{ds}s' \tag{A17}$$

Substituting the preceding expression into Equation (A13) yields:

$$\frac{1}{N}\left(W\left(s'\right)-W(0)\right)=\left(p-MSC_1(\bar{e})\right)\frac{dx_1}{ds}s'+\left(p-\frac{1}{2}s'-MSC_2\right)\frac{dx_2}{ds}s' \tag{A18}$$

The only difference between equation (1) and equation (A18) is that $MSC_1(\bar{e})$ replaces MSC_1.

A.12.2 Empirical implementation

We can use expressions (A14) and (A15) to estimate the impact on welfare in the non-linear case, and expressions (A17) and (A18) for the linear cases. We first consider the linear case and then consider an example for the non-linear case that is quadratic.

Linear MEC

Consider the case of a linear MEC function that varies with e. We follow the derivation above for expression (A15), and the formula for the MEC function we presented in appendix A.4 where we illustrate the constant MEC case.

In the constant MEC case, we evaluate the MEC function at the initial value of e(0). For the linear MEC function, following the derivation in equation (A18), we use the empirical representation of the MEC function in appendix A.4. We evaluate this function at the average congestion level before and after the subsidy, i.e., $\bar{e}=\frac{1}{2}(e'+e(0))$, where e' is the density at subsidy level s', and $e(0)$ is the density before the subsidy. The key difference between the upward sloping linear MEC case and the constant MEC case is that we allow the MEC value to vary as the subsidy is implemented.[66]

The empirical MEC function in the linear case is based on data from (Haywood and Koning, 2015). They estimate T′(e), the rate of change of the time multiplier with respect to the density (see Appendix A.4). That paper uses a contingent valuation approach to estimate how the willingness to pay for reduced congestion varies with the level of congestion. It is based on the subway system in Paris in 2010-2011, and is thus not necessarily applicable to other areas and time periods. Thus, this calculation should be viewed as illustrative.

MEC with a more flexible functional form

As a robustness check, we also implement the welfare analysis with an MEC function that is more flexible than the linear form. Specifically, we estimate the function that determines the time multiplier T(e), using the following data points taken from Haywood and Koning (2015) by applying ordinary least squares and allowing for a quadratic functional form. (Haywood

[66] We assume α_0, the constant term in the linear MEC case, is zero. That is, the marginal external cost is zero when there is no congestion.

and Koning (2015) fit a linear function, which is the value of $T'(e)$ we used for the constant and linear case). The data points for (T(e), e) are: $(1,0),(1,1),(1.05,2),(1.18,2.5),(1.26,3),(1.4,4)$ and $(1.57,6)$. The time multiplier T(e) measures how much time one is willing to travel on a crowded train with density e, relative to 1 unit of time spent on a train with no congestion (i.e., e = 0).

The slope of the fitted function gives the valuation of a marginal change in density, $T'(e) = 0.075 + 0.0106e$. This gives a new MEC function $MEC_1(e) = VOT * t * (0.075 + 0.0106e) * e + VOT * o * e$. We use the constant values described in Appendix A4 for VOT, t, and o. The resulting functional form for $\text{MEC}_1(e)$ is quadratic in e. For comparison, in the linear case, we assume $T'(e)$ is a constant.

For the quadratic MEC case, the welfare change associated with the congestion externality in the peak period is derived from equation (A13).

$$\int_0^{s'} MEC_1(e)\frac{dx_1}{ds}ds = \frac{dx_1}{ds}\int_0^{s'} MEC_1(e)ds = \frac{dx_1}{ds}\int_{e(0)}^{e'} MEC_1(e)\frac{1}{\frac{de}{ds}}de$$

where e' is the density at subsidy level s', and $e(0)$ is the density before the subsidy. We integrate the term $\int_{e'}^{e''} MEC_1(e)\frac{1}{\frac{de}{ds}}de$ numerically, assuming $\frac{de}{ds}$ to be constant for simplicity. [67]. Equivalently, as discussed in the earlier subsection, we could calculate an "average" MEC, as $\frac{1}{s'}\int_0^{s'} MEC_1(e)ds$, and substitute it in place of MEC_1 in equation (1).

Table A10 shows the results for three cases: a constant MEC function (the base case), a linear MEC function, and a quadratic MEC function. The main takeaway from Table A10 is that the welfare numbers are similar across all three cases. This reflects the facts that: 1. the MEC is only one component of welfare; and 2. The average values are not that different across all three cases (1.33 for the constant MEC case; 1.298 for the linear MEC case; and 1.1 for the quadratic MEC case).

For example, the total welfare change for the exogenous case is about \$0.029 per person per day with a constant MEC function, \$0.029 for a linear MEC function and \$0.028 for a quadratic MEC function. The results are also similar for endogenous congestion (.028,.028 and .027). Overall welfare percentage decreases in welfare (between exogenous and endogenous congestion are also similar, on the order of 5%.

It is worth pointing out that these assumptions are quite sensitive to our assumptions about the MEC function. Because there is a paucity of data on this issue, we believe more research would be useful to inform policy makers.

[67] We assume $\frac{de}{ds} = \frac{\partial e}{\partial x_1}\frac{dx_1}{ds}$ is constant, which is a reasonable approximation when the subsidy is small. See Appendix A.4 for more details.

Table A10: Results comparing welfare change with three different MEC functions

	MEC constant			MEC Linear function			MEC Quadratic function		
	Exo.	Endo.	% change	Exo.	Endo.	% change	Exo.	Endo.	% change
Welfare change associated with peak travel change	-0.0037	-0.0032	-14.7	-0.00389	-0.00332	-15.2	-0.0049	-0.0041	-15.1
Welfare change associated with off-peak travel change	0.0328	0.0309	-5.9	0.0328	0.0309	-5.9	0.0328	0.0309	-5.9
Total Welfare Change	0.0291	0.0277	-4.8	0.0289	0.0276	-4.7	0.0279	0.0267	-4.3

Notes: Numbers are calculated for the base case for the natural field experiment. Numbers may not add due to rounding. Column (1)-(2), (4)-(5) and (7)-(8) present the welfare estimates under assumption that the MEC function is constant, a linear function and a quadratic function respectively. Columns (1), (4) and (7) present welfare estimates under the assumption of exogenous congestion. Columns (2), (5), and (8) present welfare estimates under the assumption of endogenous congestion. Columns (3), (6), and (9) present the percentage change between the exogenous and endogenous cases.

A.13 Parameters for welfare analysis of the field experiment

The table below presents the parameters used for calculations for the field experiment.

Table A11: Parameters for welfare analysis of the field experiment

				VOT = 75% of wage		VOT = 50% of wage		75%	50%
Usual time (t_1)	Targetted time (t_2)	$\frac{dx_1}{ds}$	$\frac{dx_2}{ds}$	MEC1	MEC2	MEC1	MEC2	MVPF	MVPF
070000	062000	-0.02	0.01	1.13	0.30	0.05	-0.50	1.28	1.06
070000	064000	-0.02	0.01	1.15	0.58	0.06	-0.31	1.17	1.04
070000	072000	-0.02	0.00	1.51	1.90	0.31	0.56	1.13	1.03
070000	074000	-0.02	0.02	1.55	1.97	0.34	0.62	1.18	0.96
074000	070000	-0.02	0.00	1.31	0.45	0.17	-0.40	1.33	1.06
074000	072000	-0.02	0.00	1.86	1.72	0.54	0.45	1.18	1.05
074000	080000	-0.02	0.03	2.15	2.38	0.73	0.89	0.81	0.91
074000	082000	-0.02	0.01	2.15	2.71	0.73	1.11	1.37	1.11
080000	072000	-0.04	0.00	1.67	1.23	0.41	0.12	1.50	1.12
080000	074000	-0.04	0.02	1.82	1.61	0.51	0.38	1.23	1.07
080000	082000	-0.04	0.01	2.04	2.46	0.66	0.94	1.34	1.10
080000	084000	-0.04	0.03	2.07	1.91	0.68	0.57	1.48	1.18
082000	074000	-0.03	0.02	2.08	1.53	0.69	0.32	1.35	1.14
082000	080000	-0.03	0.03	2.52	2.09	0.98	0.69	1.15	1.07
082000	084000	-0.03	0.03	2.39	2.03	0.89	0.65	1.22	1.10
082000	090000	-0.03	0.01	2.01	1.06	0.64	0.01	1.72	1.28
084000	080000	-0.01	0.03	2.37	2.39	0.88	0.89	0.45	0.79
084000	082000	-0.01	0.01	1.99	2.14	0.63	0.72	1.07	1.02
084000	090000	-0.01	0.01	1.76	1.20	0.48	0.10	1.10	1.06
090000	082000	-0.02	0.01	2.01	2.82	0.64	1.18	1.19	1.04
090000	084000	-0.02	0.03	2.55	3.10	1.00	1.37	0.71	0.85
090000	094000	-0.02	0.01	1.28	0.19	0.16	-0.57	1.34	1.09

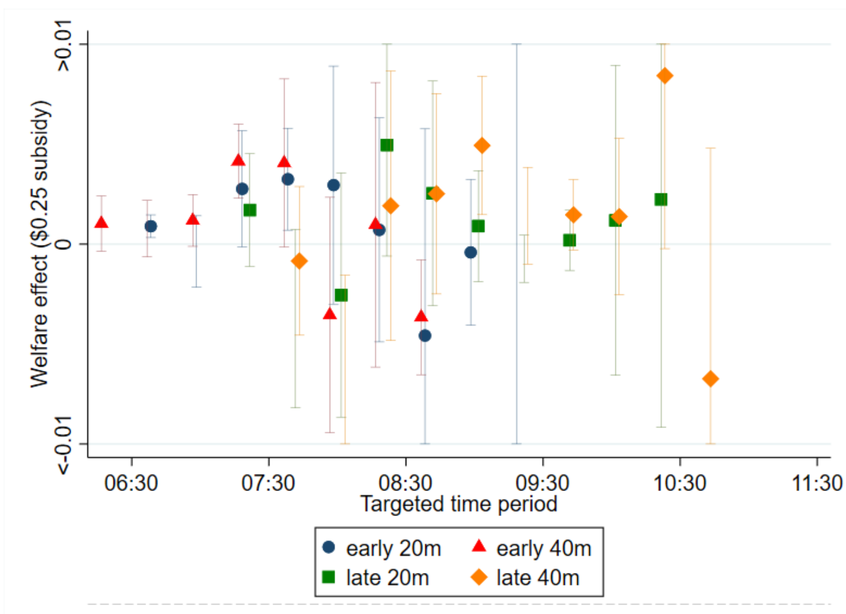
092000	084000	-0.02	0.03	1.04	1.72	0.00	0.45	0.49	0.72
092000	090000	-0.02	0.01	1.56	1.94	0.34	0.59	1.06	1.00
092000	094000	-0.02	0.01	0.93	0.28	-0.08	-0.52	1.11	1.01
092000	100000	-0.02	0.01	0.77	-0.16	-0.19	-0.81	1.32	1.12
094000	100000	-0.01	0.01	-0.02	-0.27	-0.71	-0.88	1.08	1.14
094000	102000	-0.01	0.05	0.18	-0.29	-0.58	-0.89	0.58	-0.06
100000	102000	-0.04	0.05	-0.22	-0.39	-0.85	-0.96	1.30	1.42
100000	104000	-0.04	0.01	-0.32	-0.69	-0.92	-1.16	0.92	0.69

Note: The table presents the parameter estimates for experiment 2. Column (1) indicates the usual travel time of the subsidized group, and column (2) indicates the subsidized time. Columns (3) and (4) are the demand response estimated from experiment 2. Columns (5) and (6) indicate the marginal external cost of the usual and subsidized travel time, when value of time is 75% of wage. Columns (7) and (8) indicate the marginal external cost for the usual and subsidized travel time, when the value of time is 50% of wage. Columns (9) and (10) indicate the calculated MVPF of the subsidy. The *MEC* data are based on train load data provide by BART.

A.14 Welfare estimates for the natural field experiment using the welfare model

In the body of the paper, we used MVPF as an index of welfare with different values for the value of time (VOT). We found that MVPFs could be above or below one, depending on the particular field experiment. Figure A4 shows a similar calculation using a measure of net benefits per person per weekday with a $.25 subsidy. It relies on the welfare model summarized in section 2, and uses point estimates. The key point of the figure is that some estimates for net benefits are positive and some are negative. The figure assumes that VOT is 50% of the wage. The same qualitative finding emerges in the case where VOT is 75% of the wage rate.

Figure A4: Welfare effect of the subsidy in the field experiment (VOT = 50% with sufficient statistics approach)



Note: This figure shows the welfare effect (per subsidy recipient) of a targeted subsidy. The x-axis represents the beginning time of the targeted time period. The shape of each point shows the usual travel time of the group of riders, which is 20 or 40 minutes earlier or later than the targeted time. The value of time is assumed to be 50% of the wage.

A.15 Comparing the Welfare Model with the MVPF calculation

In this section, we compare the expression for $\frac{dW(s)}{ds}$ derived using our welfare model (Equation A3) with an expression we derive for $\frac{dW(s)}{ds}$ using the MVPF approach. We will show the following lemma:

Lemma 1: If the net cost to the government is funded by a lump sum tax, then the expression for $\frac{dW(s)}{ds}$ using sufficient statistics is the same as the expression for $\frac{dW(s)}{ds}$ that we derive below for the MVPF.

Proof: A general expression for the MVPF is

$$\frac{x_2 - MEC\frac{dx_1}{ds}}{x_2 + s\frac{dx_2}{ds} - (p - MPC)(\frac{dx_1}{ds} + \frac{dx_2}{ds})} \tag{A19}$$

where we use the same notation introduced in section 2, and assume that this applies to N individuals. The numerator represents the willingness to pay and the denominator represents the net cost to the government. If we assume that this net cost is funded by a lump sum tax, then we can represent the difference between benefits and costs at the margin using the MVPF approach as the numerator minus the denominator in expression (A19). That is,

$$\begin{aligned}\frac{dW(s)}{ds} &= (x_2 - MEC\frac{dx_1}{ds}) - (x_2 + s\frac{dx_2}{ds} - (p - MPC)(\frac{dx_1}{ds} + \frac{dx_2}{ds}))\\ &= -MEC\frac{dx_1}{ds} - (s\frac{dx_2}{ds} - (p - MPC)(\frac{dx_1}{ds} + \frac{dx_2}{ds}))\\ &= (p - MPC - MEC_1)\frac{dx_1}{ds} + (p - s - MPC)\frac{dx_2}{ds} \qquad \text{(A20)}\end{aligned}$$

The second line follows from cancelling x_2 and the third from factoring. Noting that $MPC + MEC_1 = MSC_1$ and $MPC = MSC_2$ yields

$$(p - MSC_1)\frac{dx_1}{ds} + (p - s - MSC_2)\frac{dx_2}{ds}$$

This is the same expression as equation (A3) , except that equation is given in per capita terms. Multiplying equation (A3) by N gives the expression here, which proves Lemma 1.

In short we have shown conditions under which the MVPF calculation yields the same welfare result for measuring the impact of a change in the subsidy as our model that uses sufficient statistics. The general result on welfare also follows (Equation 1) if we use the same assumptions.

A.16 Comparing the demand response with exogenous congestion and endogenous congestion

In this section, we argue that the demand response to a subsidy is smaller in *absolute value* for both the peak and the off-peak periods with endogenous congestion than exogenous congestion.

Claim: The claim comes in two parts. 1. If $\frac{\partial x_1}{\partial s} < 0$ and $\frac{\partial x_1}{\partial e} < 0$, the response for peak demand in the endogenous congestion case is smaller in absolute value than that for the exogenous congestion case; and 2. In addition, if $\frac{\partial x_2}{\partial e} > 0$, the change in off-peak rides from a subsidy in the endogenous congestion case is also smaller in absolute value than that for the exogenous congestion case.

Consider the first part of the claim. We begin by computing the total derivative of peak travel with respect to a change in the subsidy. Starting with the peak demand function, $x_1(p,s,e(x_1))$, and differentiating with respect to s, yields

$$\frac{dx_1}{ds} = \frac{\partial x_1}{\partial s} + v_{11}\frac{dx_1}{ds}$$

, where $v_{11} = \frac{\partial x_1}{\partial e}\frac{\partial e}{\partial x_1} = v_{1e}\frac{\partial e}{\partial x_1}$. In particular, the total change ($\frac{dx_1}{ds}$) equals the sum of the direct price effect of the subsidy ($\frac{\partial x_1}{\partial s}$) , and the feedback effect from congestion, $v_{11}\frac{dx_1}{ds}$.
Solving for $\frac{dx_1}{ds}$, we have

$$\frac{dx_1(p,s,e))}{ds} = \frac{\frac{\partial x_1}{\partial s}}{1-v_{11}}$$

To sign v_{11}, note that because $v_{1e} < 0$ and we assume $\frac{\partial e}{\partial x_1} > 0$ (that is, increases in peak demand contribute to congestion), we have $v_{11} < 0$. This implies that

$$|\frac{dx_1}{ds}| < |\frac{\partial x_1}{\partial s}| \quad .$$

We now prove the second part of the claim. Taking the total derivative of the off-peak demand function $x_2(p,s,e(x_1))$ with respect to s gives

$$\frac{dx_2}{ds} = \frac{\partial x_2}{\partial s} + \frac{\partial x_2}{\partial e}\frac{\partial e}{\partial x_1}\frac{dx_1}{ds} \quad .$$

We use the expression for $\frac{dx_1}{ds}$ from above, which yields:

$$\frac{dx_2}{ds} = \frac{\partial x_2}{\partial s} + \frac{\partial x_2}{\partial e}\frac{\partial e}{\partial x_1}\frac{\frac{\partial x_1}{\partial s}}{1 - v_{11}} \quad .$$

Letting $v_{21} = \frac{\partial x_2}{\partial e}\frac{\partial e}{\partial x_1}$, we have

$$\frac{dx_2(p,s,e)}{ds} = \frac{\partial x_2}{\partial s} + \frac{v_{21}\frac{\partial x_1}{\partial s}}{1 - v_{11}} \quad .$$

Since $\frac{\partial x_2}{\partial e} > 0$, $v_{21} > 0$. This implies that the second term in the above equation is negative (provided that $\frac{dx_1}{ds} < 0$). This yields the result:

$$|\frac{dx_2}{ds}| < |\frac{\partial x_2}{\partial s}| \quad .$$

In other words, the increase in ridership with an off peak subsidy is less with endogenous congestion than with exogenous congestion.

These results also allow us to compare the travel in peak and off-peak periods in the endogenous congestion case with the exogenous congestion case. Let x_1^o and x_2^o be the peak and off-peak travel before the subsidy. Furthermore, let x_1' and x_2' be the peak and off-peak travel after the subsidy with exogenous congestion; and $\tilde{x}_1, \tilde{x}_2$ be the peak and off-peak travel after the subsidy with endogenous congestion case. Noting $x_1' = x_1^0 + \frac{\partial x_1}{\partial s}s'$ and $\tilde{x}_1' = \tilde{x}_1^0 + \frac{dx_1}{ds}s'$, we have $x_1' < \tilde{x}_1$ because $\frac{dx_1}{ds} > \frac{\partial x_1}{\partial s}$. Using a similar argument, $x_2' > \tilde{x_2}$ because $\frac{dx_2}{ds} < \frac{\partial x_2}{\partial s}$.

A.17 Welfare analysis of Perks subsidy with different demand responses

Riders who join the Perks program may have different demand responses than the BART riders who do not. We provide a theoretical model below that allows for different demand responses for these two groups in estimating the welfare gains of a subsidy. Allowing for different price responses enables us to put bounds on what the welfare effects could be as we scale these price subsidies up to the entire BART ridership.

Consider two types of riders, one who joins the Perks program (type a) and one who does not (type b). There are n_a type a riders and n_b type b riders, exogenously given. Define the indirect utility function for each type as $V_a(p,s,e)$ and $V_b(p,s,e)$ respectively, and define the "average" consumption, x_1 and x_2 as,

$$x_k = \frac{1}{N}(n_a x_k^a + n_b x_k^b) = \mu_a x_k^a + \mu_b x_k^b$$

for $k \in \{1,2\}$, where $N = n_a + n_b$, x_1^a and x_2^a are peak and off-peak consumption for type a consumers, and x_1^b and x_2^b are similarly defined for type b consumers. $\mu_j = \frac{n_j}{N}$ is the share of population of type j, for $j \in \{a,b\}$.

We consider the welfare impact of a single uniform subsidy for both type of riders. Welfare is given by the utility of type a, plus the utility of type b, minus the transfer to BART and the subsidy cost. The formula is

$$W(s) = n_a V_a(p,s,e) + n_b V_b(p,s,e) - \min C(Nx_1, Nx_2) - Np(x_1 + x_2) - n_a x_1^a s - n_b x_1^b s$$

with the indirect utility function given by $V_j(p,s,e) = \max_{x_1^j, x_2^j} u(x_1^j, x_2^j, e) + Z - p x_1^j - (p-s)x_2^j$ for $j \in \{a,b\}$.

The marginal welfare change from a change in s is

$$\frac{dW}{ds} = (\sum_{j=a,b} n_j (x_2^j + \frac{\partial V_j}{\partial e}\frac{de}{ds} - n_j \frac{dx_2^j}{ds} s - n_j x_2^j)) + N(p-c)(\frac{dx_1}{ds} + \frac{dx_2}{ds})$$

where c is the marginal cost of the cost function $c = C_1(Nx_1, Nx_2) = C_2(Nx_1, Nx_2) = MPC$ (and C_1 and C_2 are partial derivatives with respect to the first and second arguments in the cost function).

This expression simplifies to

$$\frac{dW}{ds} = n_a(\frac{\partial u_a}{\partial e}\frac{de}{ds}) + n_b(\frac{\partial u_b}{\partial e}\frac{de}{ds}) - n_a \frac{dx_2^a}{ds} s - n_b \frac{dx_2^b}{ds} s + N(p-c)(\frac{dx_1}{ds} + \frac{dx_2}{ds}), \tag{A21}$$

noting that $\frac{\partial V_a}{\partial e} = \frac{\partial u_a}{\partial e}$ and $\frac{\partial V_b}{\partial e} = \frac{\partial u_b}{\partial e}$ at the optimum (by the envelope theorem).

Congestion is determined by the technology $E(.)$

$$e = E(Nx_1) = E(n_a x_1^a + n_b x_1^b)$$

The impact of the subsidy on congestion is given by

$$\frac{de}{ds} = E'(Nx_1)(n_a \frac{dx_1^a}{ds} + n_b)\frac{dx_1^2}{ds} = \frac{\partial e}{\partial x_1}\frac{1}{N}(n_a \frac{dx_1^a}{ds} + n_b \frac{dx_1^b}{ds}) = \frac{\partial e}{\partial x_1}(\mu_a \frac{dx_1^a}{ds} + \mu_b \frac{dx_1^b}{ds}) \tag{A22}$$

.

Define $MEC_1 = -(\mu_a \frac{\partial u_a}{\partial e} + \mu_b \frac{\partial u_b}{\partial e})\frac{\partial e}{\partial x_1}$. It represents the the marginal external cost of an increase in total peak rides, as $(\mu_a \frac{\partial u_a}{\partial e} + \mu_b \frac{\partial u_b}{\partial e})\frac{\partial e}{\partial x_1} = (n_a \frac{\partial u_a}{\partial e} + n_b \frac{\partial u_b}{\partial e})\frac{\partial e}{\partial N x_1}$.

Substituting (A22) into Equation (A21) yields:

$$\frac{dW}{ds} = n_a(p-c-MEC_1)\frac{dx_1^a}{ds} + n_a(p-c-s)\frac{dx_2^a}{ds} + n_b(p-c-MEC_1)(\frac{dx_1^b}{ds}) + n_b(p-c-s)\frac{dx_2^b}{ds}. \tag{A23}$$

The first two terms represent the change in welfare associated with a change in type a demand in the peak and off-peak markets, and the second two terms represent the change in welfare associated with a change in type b demand in the peak and off-peak markets. Thus, the average per capita welfare change from a change depends on the demand characteristics of both groups and their respective sizes.

Equation (A23) can be seen as a generalization of 1 in the text, which measures the welfare change for one group that received the subsidy. Here, we are measuring the welfare change for two groups

Exogenous congestion when other riders (type b) are less responsive to prices by some fraction

We can simplify equation (A23) if we assume that the demand response in both markets is similar for those consumers who do not join Perks. Specifically, assume that type b riders have a demand response that is proportional to type a in both markets. Specifically, we let $\frac{dx_j^b}{ds} = k\frac{dx_j^a}{ds}$ where k is between 0 and 1, for $j \in \{1,2\}$. Substituting in equation (A23) yields:

$$\frac{dW}{ds} = (n_a + n_b k)(p-c-MEC_1)\frac{dx_1^a}{ds} + (n_a + n_b k)(p-c-s)\frac{dx_2^a}{ds}$$

Integrating this expression between 0 and s' yields the total welfare change:

$$W(s') - W(0) = \int_0^{s'} \frac{dW}{ds} ds = (n_a + n_b k)(p - c - MEC_1)\frac{dx_1^a}{ds} s' + (n_a + n_b k)(p - c - \frac{1}{2}s')\frac{dx_2^a}{ds} s'$$

The welfare change per N population is

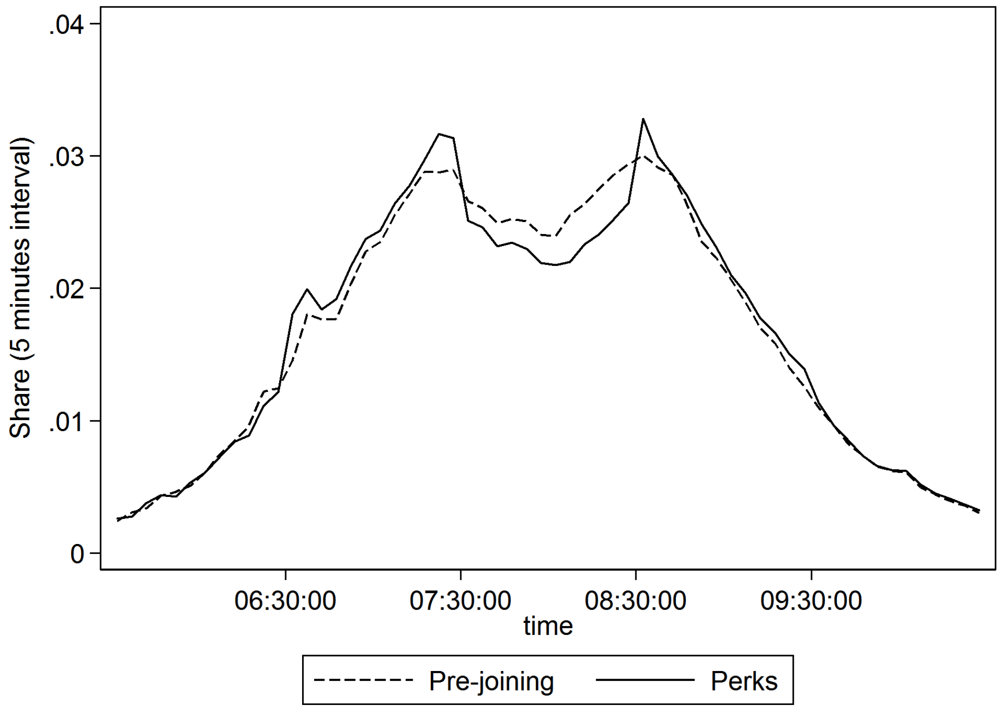
$$\begin{aligned} \frac{1}{N}(W(s') - W(0)) &= \frac{(n_a + n_b k)}{N}(p - c - MEC_1)\frac{dx_1^a}{ds} s' + \frac{(n_a + n_b k)}{N}(p - c - \frac{1}{2}s')\frac{dx_2^a}{ds} s' \\ &= \mu((p - c - MEC_1)\frac{dx_1^a}{ds} s' + (p - c - \frac{1}{2}s')\frac{dx_2^a}{ds} s') \end{aligned} \quad \text{(A24)}$$

where $\mu = \frac{n_a + n_b k}{N}$ is between 0 and 1.

The formula here for two groups is very similar to the formula for welfare change for one group in the text (see equation 1), with the difference being we have added the factor μ. Note that when the demand responses of the two groups are identical (k=1), the two formulas are the same.

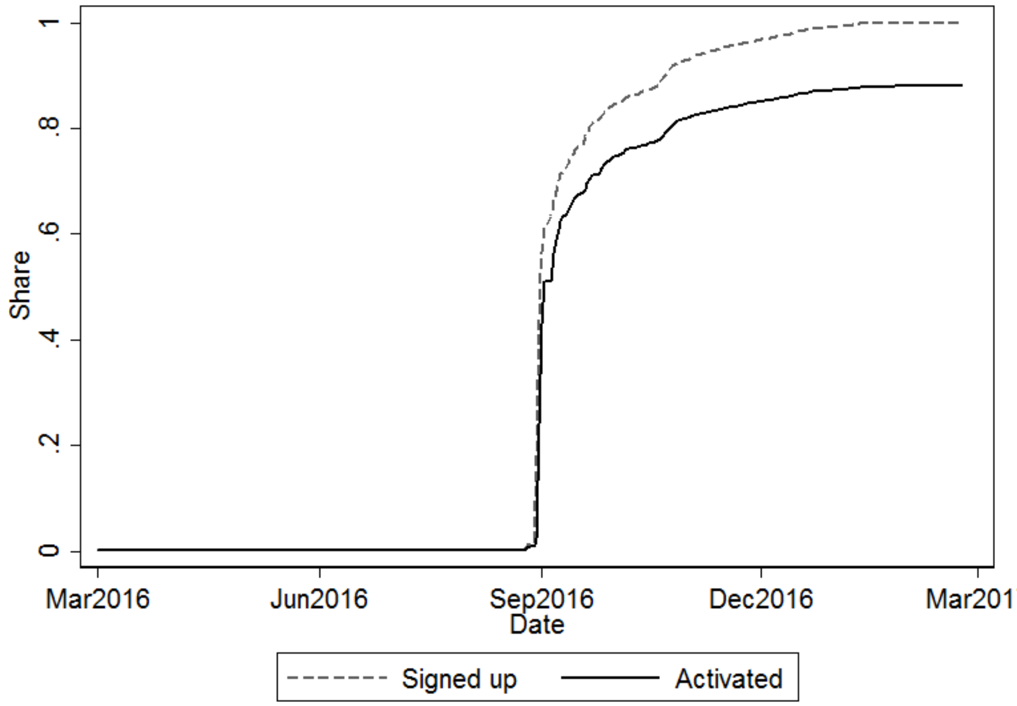
B Additional Figures and Tables

Figure A5: Natural experiment - Travelling time of participants



Note: The figure plot the density of trip by its starting time for participants in the first experiment. Sample includes trips taken between with start time between 5.30am to 10.30am on Mon-Fri. Sample period include Mar 2016 to Feb 2017, that includes 2,099,895 trips by 17,788 participants. Bin width is 5 minutes interval.

Figure A6: Program enrollment



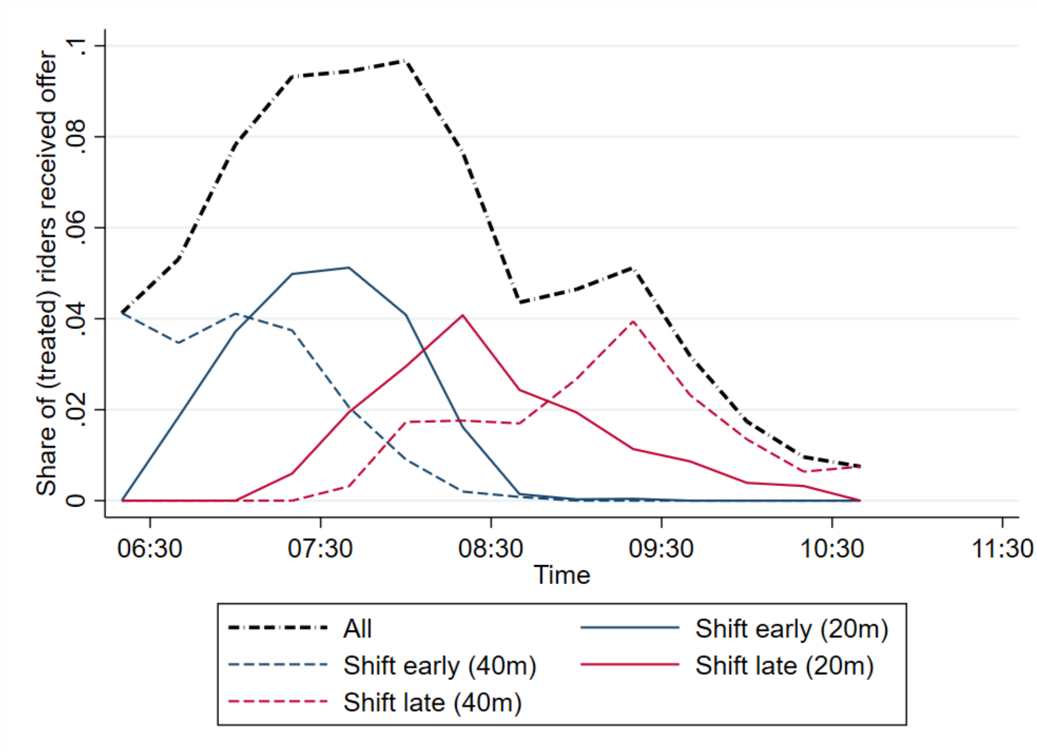
Note: The graph plots the cumulative share of users signed up, and who had activated.

Table A12: Natural experiment - Long term effect DiD

	Outcome: Share of all trip in morning	
	All dates	
	Peak Hour	Bonus Hour
	(1)	(2)
Treated users × Perks period	-0.0211***	0.0255***
	(0.00133)	(0.000963)
Treated users × Post-Perks period	-0.00946***	0.00135
	(0.000909)	(0.000927)
Treated users	-0.00297***	0.0744***
	(0.000554)	(0.000644)
Observations	636	636

Note: Daily observation for Perks users and the rest of network. Sample includes 1 Apr 2016-30 Jun 2017, weekdays excluding public holidays. Robust standard error reported in parenthesis.

Figure A7: Proportion of actual offers given (a_j)



Note: This figure plot the share of user-date under each of the types of offers, during the program period among the treated (in (a)) users.

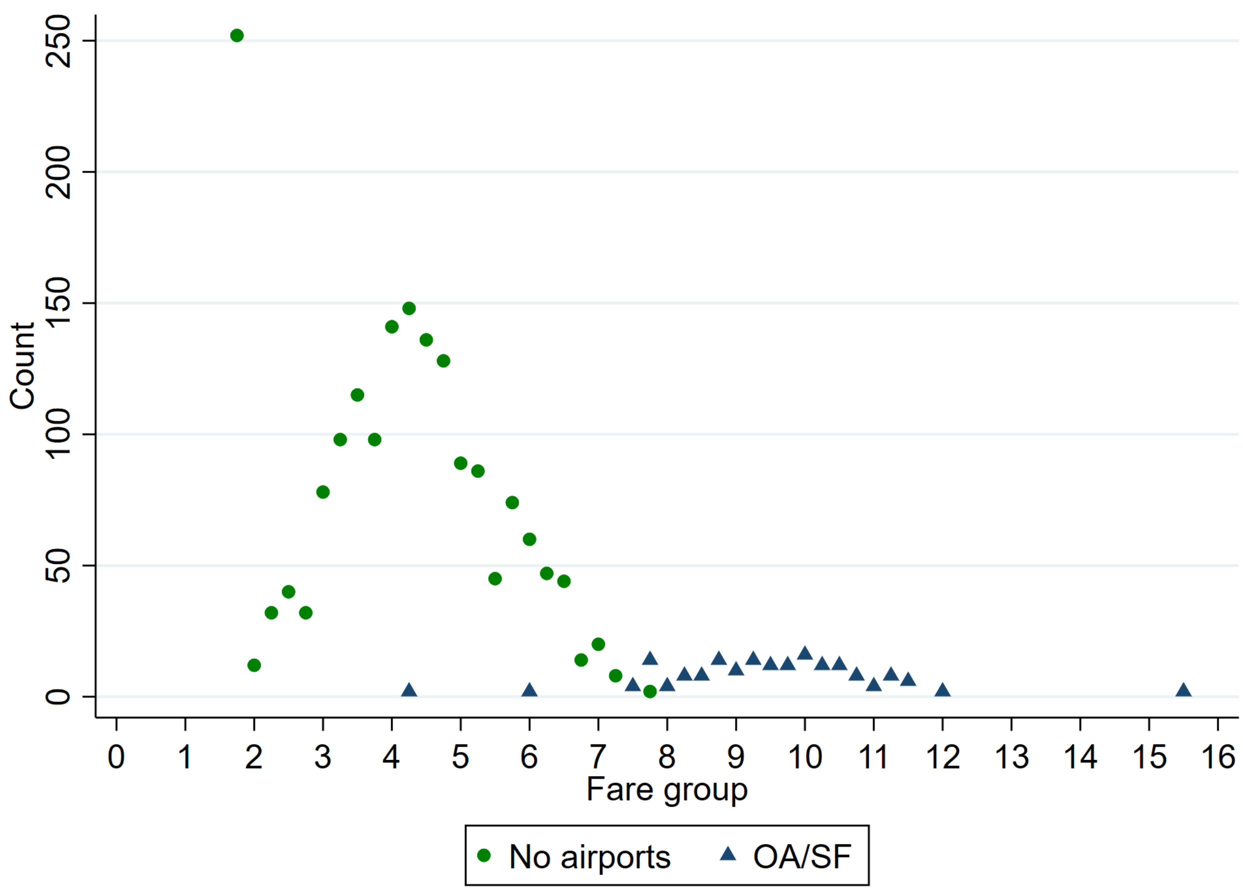
Table A13: Route level difference-in-differences estimate

	Outcome: Share of all trip in morning	
	Peak Hour	off-peak shoulder hour
	(1)	(2)
Treated users × Perks Period	-0.0332***	0.0225***
	(0.00428)	(0.00416)
Observations	97457	95432
Route-Perks user FE	x	x
Week FE	x	x

Notes: The table reports estimation of equation 2 at route level at weekly level. The outcome is weekly observation for each route in the BART network for Perks users or the rest of network. Sample includes 1 Aprl 2016 to 28 Feb 2017, all weekdays excluding specific public holidays. Standard error clustered at route-group level.

Figure A8: BART system map

Figure A9: Distribution of fare for each route on BART



Notes: The figure plots the distribution of fare for each route on BART. The diamond markers represent the fare distribution for routes excluding those to or from the San Francisco International Airport or Oakland International Airport. The triangle markers represent the distribution for routes that are to or from the San Francisco International Airport or Oakland International Airport. (2017 prices)

Figure A10: Example of ads to motivate sign-up in the natural experiment

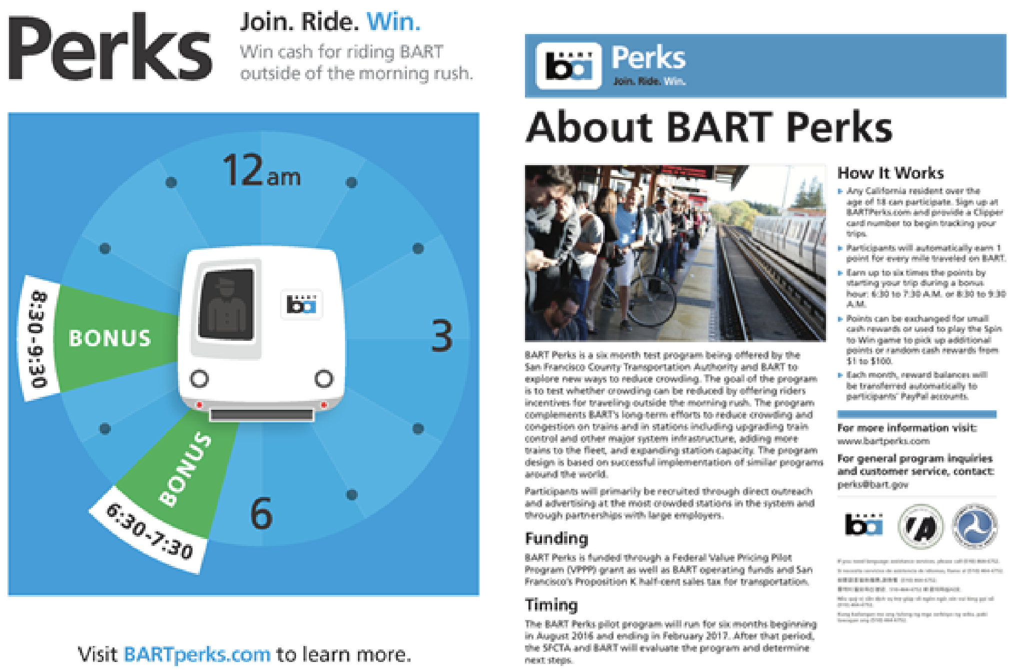
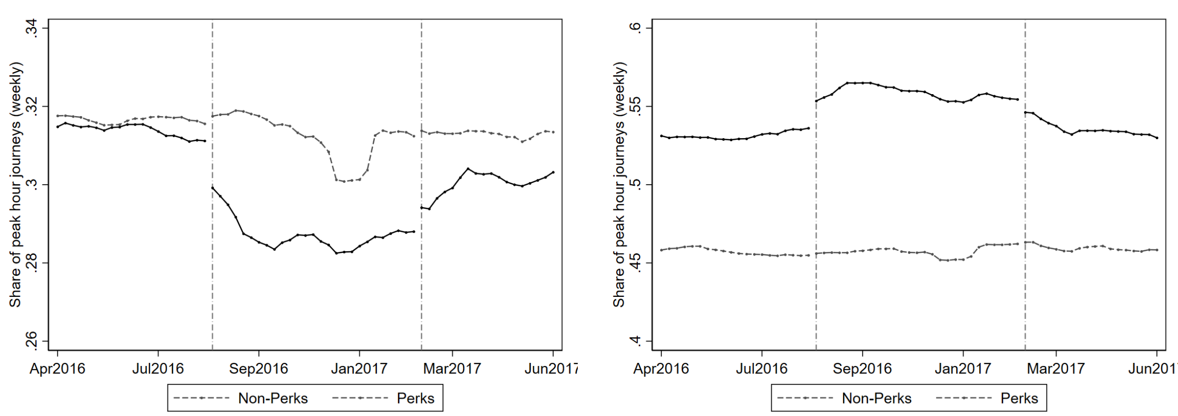


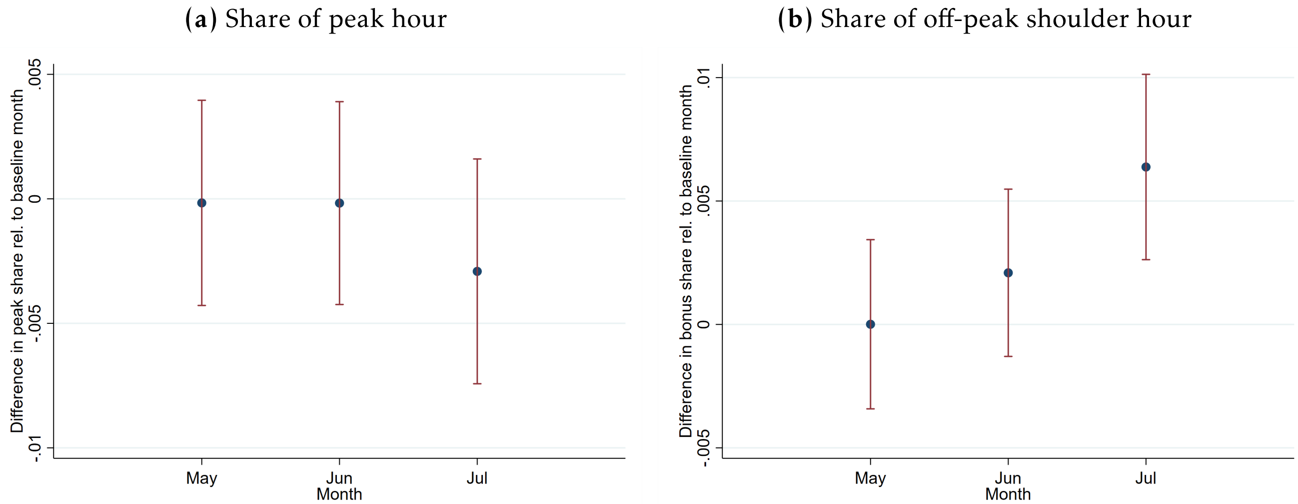
Figure A11: Long term - Share of peak hour journeys

(a) Peak hour (b) Off-peak shoulder hour



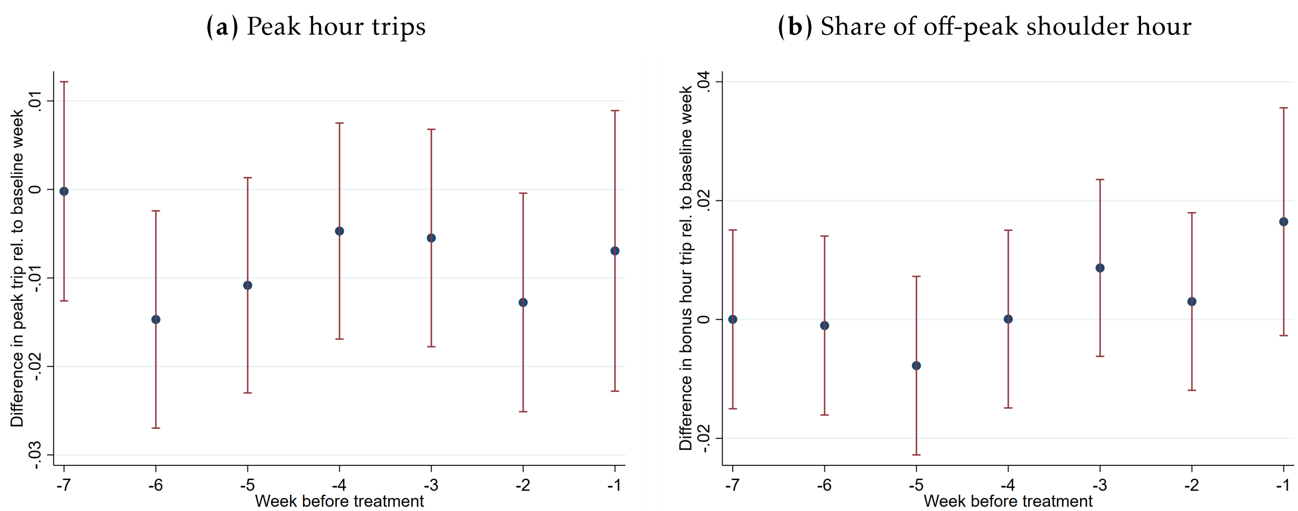
The graph plot trend of weekly average peak hour (panel (a)) and off-peak shoulder hour (panel (b)) journey share in the morning (6:30am-7:30am and 8:30am-9:30am among all journeys in 5:30am-10am), for weekdays. Sample include 1 Apr 2016-30 Jun 2017, excluding public holidays.

Figure A12: Test of pre-trend



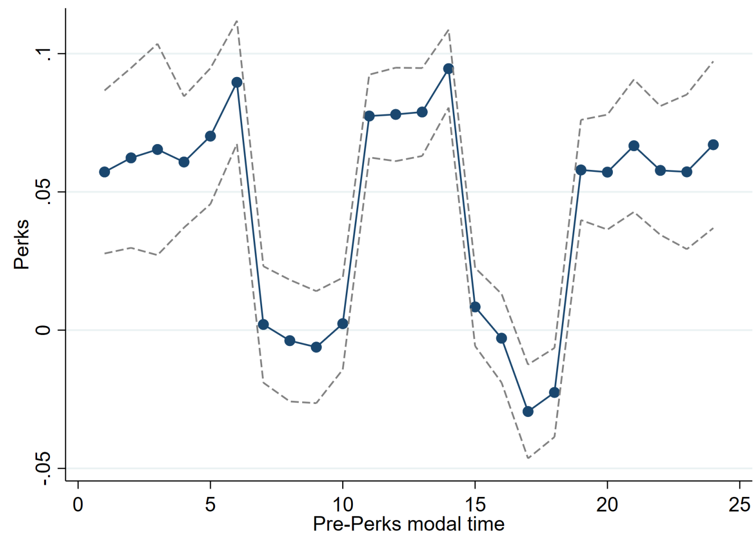
The figure plots the test of pre-trend with respect to Figure 1 and 2. It estimate a variant of equation 2, which is $y = \beta_1 * Treated * May + \beta_2 * Treated * June + \beta_3 * Treated * Jul + \eta_g + \gamma_t + \epsilon_{gt}$, where y_{gt} is the share of peak hour trips (panel (a)) or share of off-peak shoulder hour trips (panel (b)) in the morning from 5.30am-10.30am on date t for group g. May, June and Jul are indicators for month May, June and July 2016 respectively. Sample includes dates from April 2016 to 23 Jul 2016. β_1, β_2 and β_3 represent respectively the change of difference between the Perks participants and the rest of the network, from April to the respective month.

Figure A13: Test of pre-trend



The figure plots the test of pre-trend with respect to Figure 1 and 2. It estimate a variant of equation 3, which is $y_{it} = \sum_k \beta_k * Treated * D_k + \sigma_i + \mu_t + \epsilon_{it}$, where y_{it} is the number of trip during peak hour (panel (a)) or bonus hour (panel (b)). D_k is an indicator of each week before the treatment in 23 Aug. Sample includes dates from 2016 week 27 to week 34. β_k represent respectively the change of difference between the early enrolled participants and the late enrolled participants, from week 27 to the respective week of the year. Baseline/omitted week is 2016 week 27.

Figure A14: Implied treatment effect on off-peak shoulder hour travel by modal travel time in pre-treatment period



Notes: The figure plots the treatment effect on off-peak shoulder hour travel implied by the estimation of equation (4) with additional interaction of the term $PerksActive_{it}$ with indicators of the modal travel time of user i in period before perks. The modal travel time are defined by 15 minutes intervals.

Figure A15: Example of ads to motivate sign-up in the field experiment

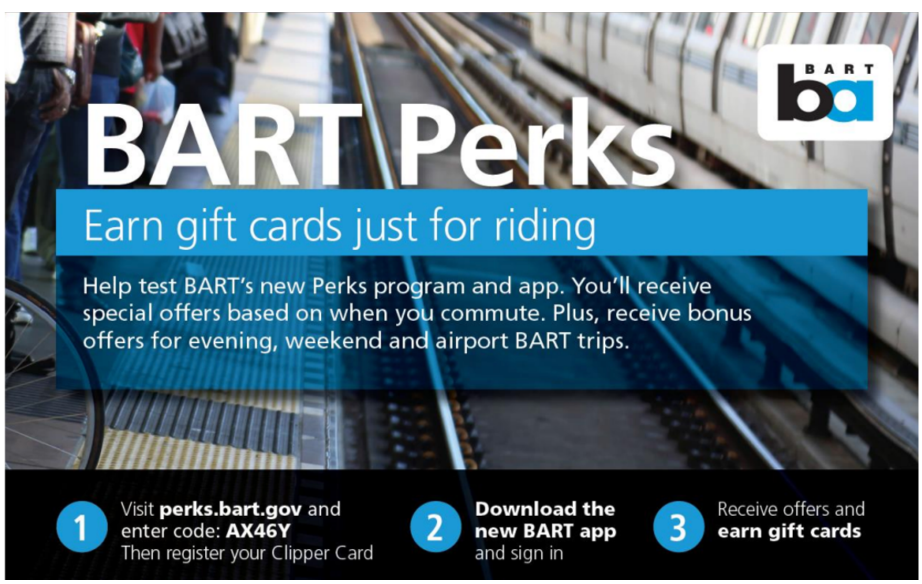
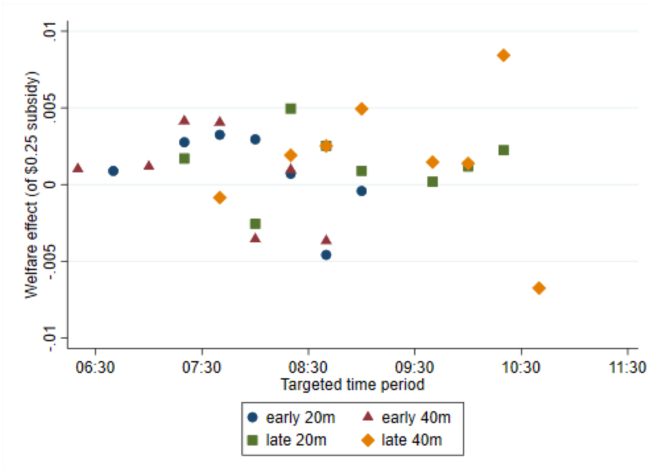
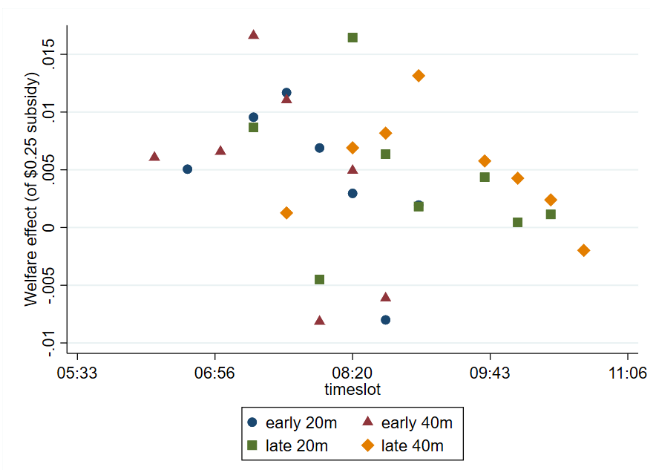


Figure A16: Welfare effect of subsidy in the field experiment – by subsidized time

(a) VOT - 50% of wages



(b) VOT - 75% of wages



Note: This figure shows that welfare change of a targeted subsidy can be positive or negative. The figure plots the welfare effect of the subsidy per rider using equation 1". The x-axis represents the beginning time of the targeted time period. The shape of each point shows the usual travel time of the group of riders, which is 20 or 40 minutes earlier or later than the targeted time. The value of time is assumed to be 50% of the wage.

Table A14: Average treatment effects by time period: Early vs. late comparison with alternative definition

	OLS (late enrolled users: Oct 15 to Nov 5)		
	Outcome: Number of trips daily		
	Peak	Off-peak shoulder (Either)	Total (5.30-10.30am)
	(1)	(2)	(3)
Early enrolled users × Perks Period (β_1)	-0.00817**	0.0125***	0.00105
	(0.00355)	(0.00428)	(0.00497)
Observations	582017	582017	582017
Date FE	x	x	x
User FE	x	x	x
Sample mean (before Perks)	0.18	0.31	0.57

Note: The sample includes daily trips from June 1 to October 15, 2016, excluding August 23 to September 2. The Perks period is defined as dates after September 2. Early enrollers are defined as those who enrolled between August 23 and September 2. The control group includes those who enrolled between October 15 and November 5, 2016. Standard errors are clustered at user-week level. The sample means are for late enrolled users in periods before the Perks period.

Table A15: Route level difference-in-differences estimate - by direction of travel towards Montgomery Street and Embarcadero

	Outcome: Share of all trip in morning			
	Peak hour		off-peak shoulder hour	
	(1)	(2)	(3)	(4)
Treated useres × Perks period	-0.0213***	-0.00323	0.0216***	-0.00108
	(0.00485)	(0.00707)	(0.00531)	(0.00744)
N	4793	1416	4791	1421
Route-Perks user FE	x	x	x	x
Week FE	x	x	x	x
Sample	Westbound	Eastbound	Westbound	Eastbound

Notes: The table reports estimation of equation 2 at route level at weekly level. The outcome is weekly observation for each route in the BART network for Perks users or the rest of network. Sample includes 1 April 2016 to 28 Feb 2017, all weekdays excluding specific public holidays. Standard error clustered at route-group level. Sample includes routes exiting at Montgomery Street and Embarcadero. Column (1) and (3) report the estimate for the sample with origin travelling westbound towards Montgomery Street and Embarcadero. Column (2) and (4) report the estimate for the sample with origin travelling eastbound towards Montgomery Street and Embarcadero.

Table A16: Treatment effects on early or late off-peak shoulder hour: staggered activation

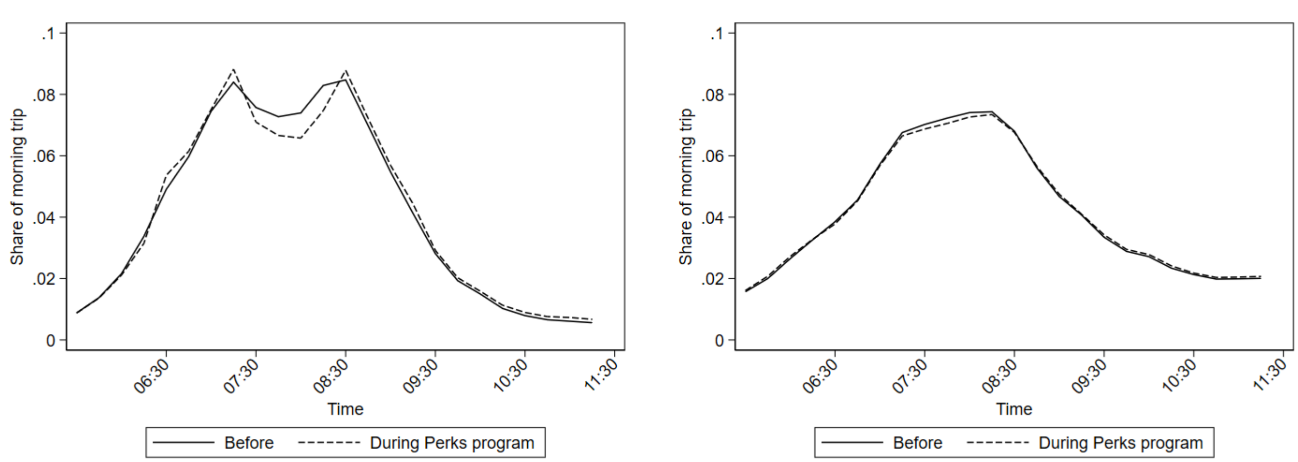
	OLS	
	Outcome: Number of daily trips	
	Off-peak shoulder hour (early)	Off-peak shoulder hour (late)
	(1)	(2)
Active status level	0.0175***	0.00992***
	(0.00261)	(0.00289)
Observations	821643	821643
Date FE	x	x
User FE	x	x
Baseline mean	0.291	0.254
Implied own-price elasticity	-0.653	-0.424

Note: Sample includes dates from 1 Jul-6 Nov 2016. Standard errors clustered by users and dates. Early off-peak shoulder hour refers to trip between 6.30-7.30am; late off-peak shoulder hour refers to trip between 8.30-9.30am.

Figure A17: Comparing Perks program participants with the rest of the BART network

(**a**) Perks participants

(**b**) Rest of the BART network



Note: The figure plots the average daily share of trips taken by the participants in the first experiment in each time interval from 530am-1130am. It also plots on the same figure the average share of daily trips at each time interval taken by the rest of the BART users. The sample for Before Perks period includes April 1, 2016 to August 22, 2017 for all weekdays. The sample for Perks peiod includes August 23, 2016 to February 28, 2017 for all weekdays.

Table A17: MVPF - average and transbay

	All		Transbay	
	Exo. congestion	Endo	Exo.	Endo
MVPF	1.598 [1.479, 1.716]	1.563 [1.444, 1.681]	1.6818 [1.5602, 1.803]	1.6336 [1.515, 1.751]

Table A18: Alternative DiD estimator - staggered activation for the natural experiment

	Outcome: Number of trips daily			
	Peak	Off-peak shoulder (Either)	Other	Total
Treatment effect (Instantaneous)	-.0158	.00540	-.00495	-.0153
(De Chaisemartin and d'Haultfoeuille, 2020)	(0.0038)	(0.00513)	(0.00317)	(0.00523)
ATT (weekly data)	-0.0137	0.00590	-0.0074	-0.0155
	(0.004)	(0.0052)	(0.0029)	(0.0054)
(Callaway and Sant'Anna, 2021)	(0.0046)	(0.0055)	(0.0034)	(0.0056)
Observations	1426285	1426285	1426285	1426285

Note: The table present the alternative DiD estimator for the treatment effect of subsidy in experiment 1. The first row report the estimate of the instantaneous treatment effect of activation in experiment 1 following (De Chaisemartin and d'Haultfoeuille, 2020). Standard errors are calculated with 50 bootstrap replication. The second row report the average treatment effect following Callaway and Sant'Anna (2021) estimated with daily data averaged at weekly level, using the "not yet treated" as control group.

Table A19: Summary statistics - field experiment

	Period before experiment	Experiment period	
	Average no. of trip by time period	Share of user-date received experiment offer	
		Shift early offers	Shift late offer
Time period	mean	mean	
062000	0.020	0.022	
064000	0.028	0.028	
070000	0.064	0.041	
072000	0.100	0.045	0.003
074000	0.106	0.037	0.012
080000	0.107	0.026	0.024
082000	0.083	0.009	0.030
084000	0.054	0.001	0.021
090000	0.030	0.000	0.024
092000	0.017	0.000	0.026
094000	0.012		0.017
100000	0.008		0.009
102000	0.005		0.005
104000	0.003		0.004
N	196083	96982	96982

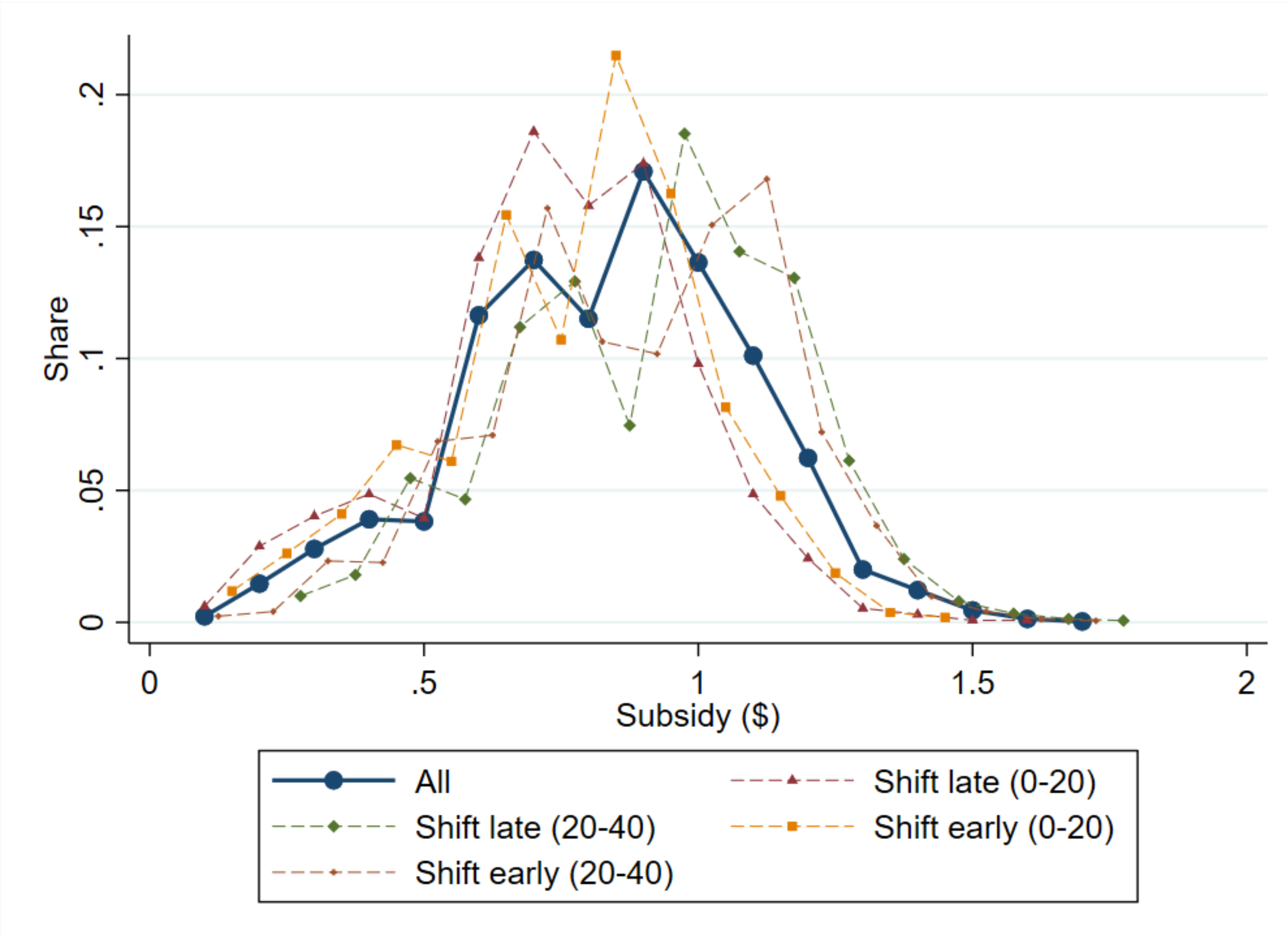
Notes: The table report summary statistics for experiment 2. Column (1) report the participants average daily number of trip by 20 minutes time interval in period before the experiment. Column (2) and (3) report on average for each date during the experiment period, the number of user received offer by the incentivized time period in the morning. Column (2) (and (3)) report the average number of users receive an offer that the incentivized offer is earlier (later) than their usual travel time.

Table A20: Journey characteristics

	(1) All BART	(2) Perks pre	(3) Early pre	(4) Late pre
Distance (mile)	14.60	15.22	15.07	13.77
	(9.05)	(9.76)	(9.48)	(9.33)
Fare ($)	3.89	4.05	4.04	3.90
	(1.31)	(1.30)	(1.29)	(1.26)
Duration (s)	NA	2226.83	2216.75	2071.28
	NA	(1094.26)	(1073.71)	(1082.17)
N		23,770	13,481	2,225

Notes: The table report the average journey characteristics for all BART morning trips, and for the Perks sample. Column (1) report the journey characteristics for all BART trips between Jun-Aug 2016. Columns (2) present the journey characteristics for the Perks sample for the period between Jun-Aug 2016. Column (3) and (4) present the characteristics for the Early and late enrolled perks riders, respectively during Jun-Aug 2016.

Figure A18: Subsidy in experiment



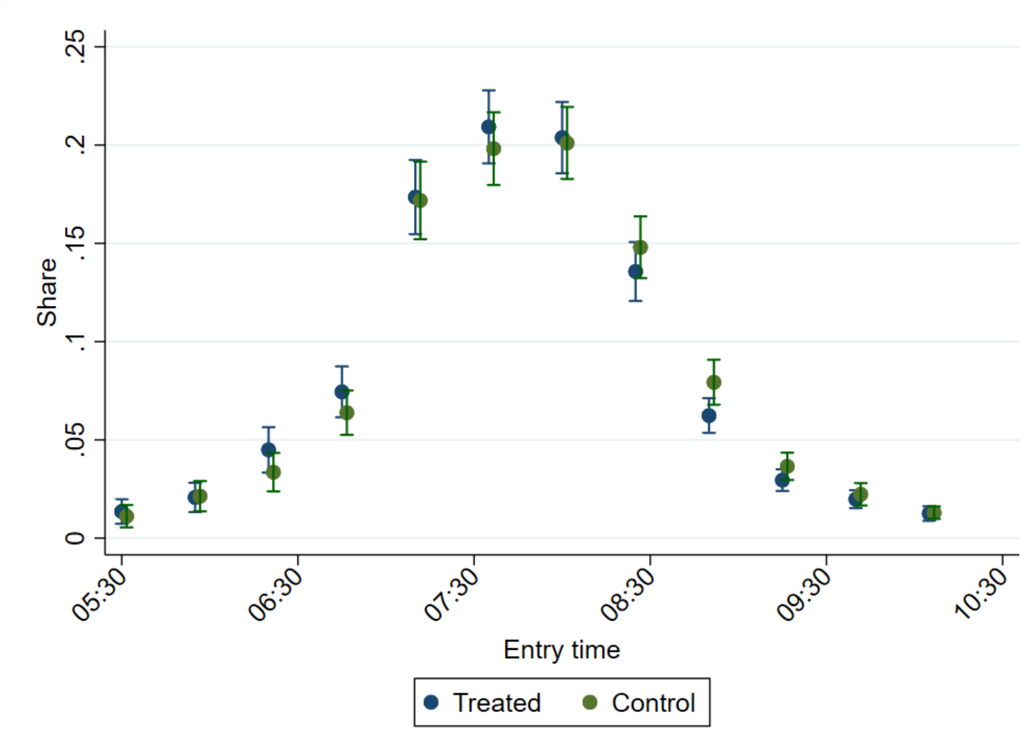
Notes: The figure plots the distribution of subsidy amount in experiment 2 for the morning subsidies. The y-axis represent the share of subsidy given with the amount (in bin width of 0.1$), among all subsidies in the respective type.

Table A21: Decomposition of the staggering enrollment specification

	Rush Hour	Bonus Hour	Total	(Weight)
Overall	-0.007	0.024	0.014	
	(0.001)	(.0016969)	(0.002)	
Timing groups	-0.008	0.025	0.014	0.783
Never v timing	-0.003	0.021	0.014	0.217

Notes: the table present the Goodman-Bacon decomposition for the staggering enrollment specification. The first row present the overall DiD estimates on a fully balanced sub-sample. The second and third row present the estimates by decomposition. The fourth column present the weights.

Figure A19: Comparison of treated and control group before experiment



Notes: The figure plots the share of entry time for the trip among all trips between 5.30am-10.30am, for the treated and control group respectively before the experiment. It also plot the 95% confidence interval for the share clustering at user level.

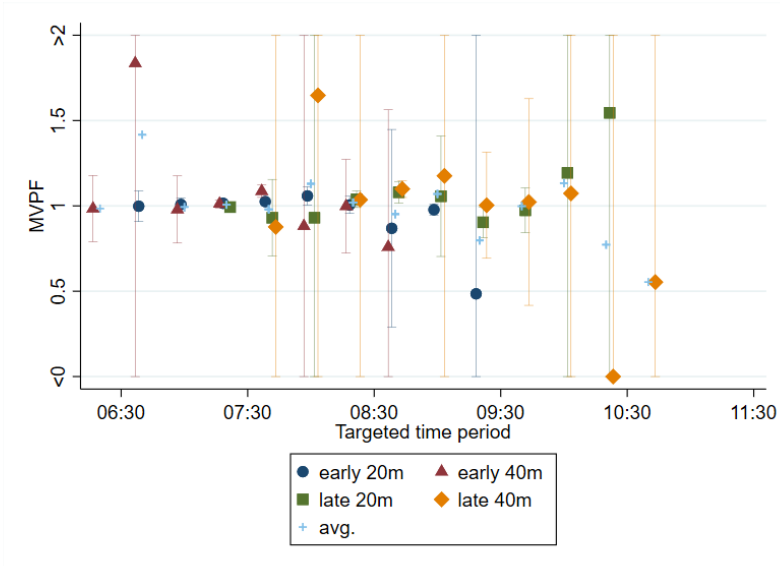
Table A22: Comparison of treatment and control group for experiment

	(1) Distance	(2) Fare
Treated	1155.9	0.0771
	(737.3)	(0.0648)
N	145066	149933
Mean	26459.42	4.40

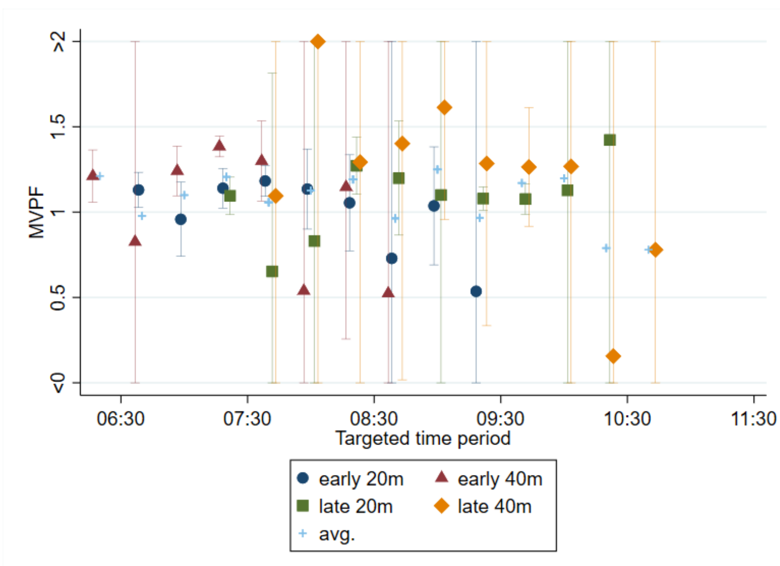
Notes: The table presents a regression of the outcomes on a treatment indicator for the field experiment, for the period three months before the experiment.

Figure A20: MVPF of subsidy in experiment 2 - including mode shifting

(a) VOT - 50% of wages



(b) VOT - 75% of wages



Note: This figure shows the MVPF of a targeted subsidy. See text for details. The x-axis represents the beginning time of the targeted time period. The shape of each point shows the usual travel time of the group of riders, which is 20 or 40 minutes earlier or later than the targeted time. The value of time is assumed to be 50% of the wage in panel (a) and 75% of the wage in panel (b).

Table A23: Average effect of experiment 2 subsidies - by shift early and shift late subsidies

	Usual travel time		Subsidized time	
	(1)	(2)	(3)	(4)
$T^{Shiftearly}$	-0.0353**		0.0336***	
	(0.0163)		(0.00928)	
$T^{Shiftlate}$		-0.0458***		0.0217***
		(0.0171)		(0.00820)
N	31501	24851	31501	24851

Note: The table present the estimates for equation 5 and 6. The outcome variables for column (1)-(2) is an indicator for rider i who received an shift early or shift late offer on date t and travelled at the subsidized time. The outcome variables for column (3)-(4) is an indicator for rider i who received an subsidies (actual or hypothetical) on date t and travel on the usual travel time, corresponding to that of the subsidy offer given. Sample in column (1) and (3) includes treated and control riders for the dates they received a shift early subsidies (actual or hypothetical). Sample in column (2) and (4) includes treated and control riders for the dates they received a shift late subsidies (actual or hypothetical). Standard errors are clustered at rider level.