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.

Appendix A.1 – Welfare model

This section presents the derivation of the key welfare equation that is used in the theoretical model.

Appendix A.2 - Optimal level of the subsidy

This section presents the derivation of the optimal subsidy in the theoretical model.

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

This section presents an extension of the theoretical model in section 2 with multiple markets and congestion levels.

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

This section presents the methodology and empirical estimates that we used to estimate the marginal external cost of a trip.

Appendix A.5 - Estimation of the marginal cost per trip

This section presents the estimation of marginal private cost for a BART trip that we used in the welfare analysis in section 5.

Appendix A.6 - Estimation of feedback effect from congestion

This section presents the methodology and empirical estimates on the effect of congestion on travel demand for the welfare analysis in section 5.

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

This section presents the additional parameters of the welfare calculation of the natural experiment 1 that we used in section 5.

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

This section presents an additional table on the welfare effect of a subsidy in the natural field experiment, including both the Transbay route and non-Transbay routes.

Appendix A.9 - Welfare and congestion impacts for the optimal subsidy

This section presents the optimal subsidy for the natural experiment and the welfare effects

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:

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

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:

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

[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:

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

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.

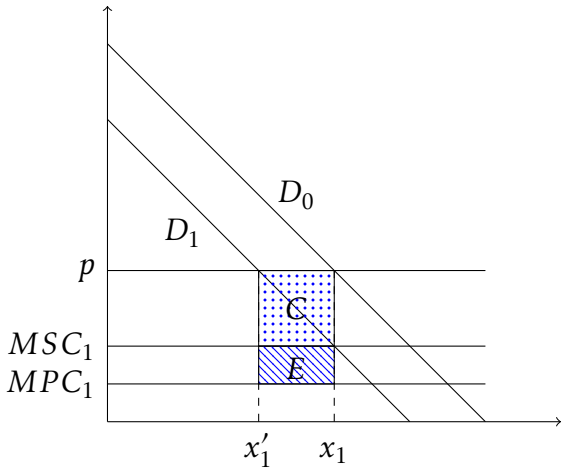
The graph for the off-peak market is shown in Figure A2. The off-peak demand is given by D. The subsidy price is reduced from p to $p - s$, with an increase in demand from x_2 to x_2'. Because there is no change in congestion in the off-peak period (it is presumed to be zero), we

[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.

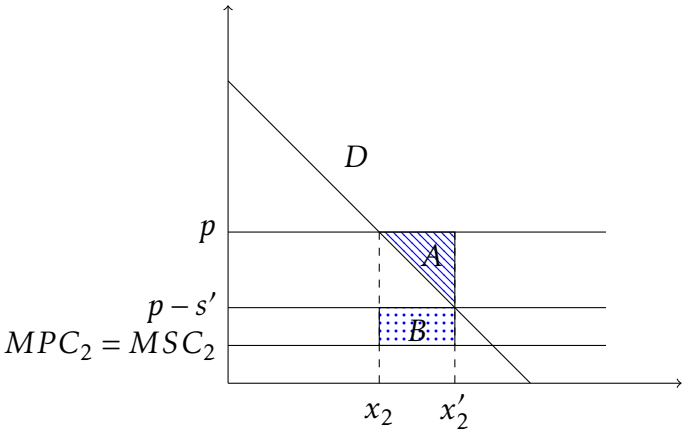
[54] We ignore any direct changes in consumer utility resulting from the subsidy in Figure A1 and Figure A2 below because of the envelope theorem.

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 gradu-

ally 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:

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

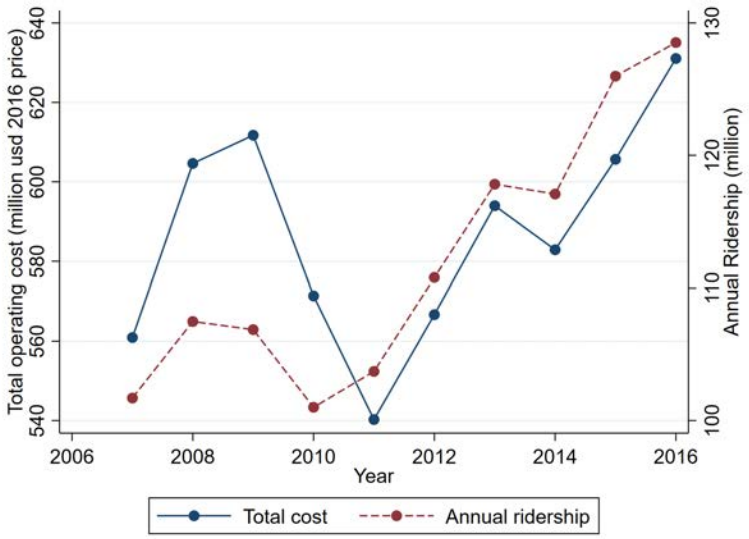
where $C(Nx)_t$ is the total operating cost of BART in year t, Nx_t is the annual number of total trip in the system, and β represent the marginal private cost per BART trip.

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.

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

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

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

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., $\bar{e}=\frac{1}{2}\left(e'+e(0)\right)$, 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 simplicitly. [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

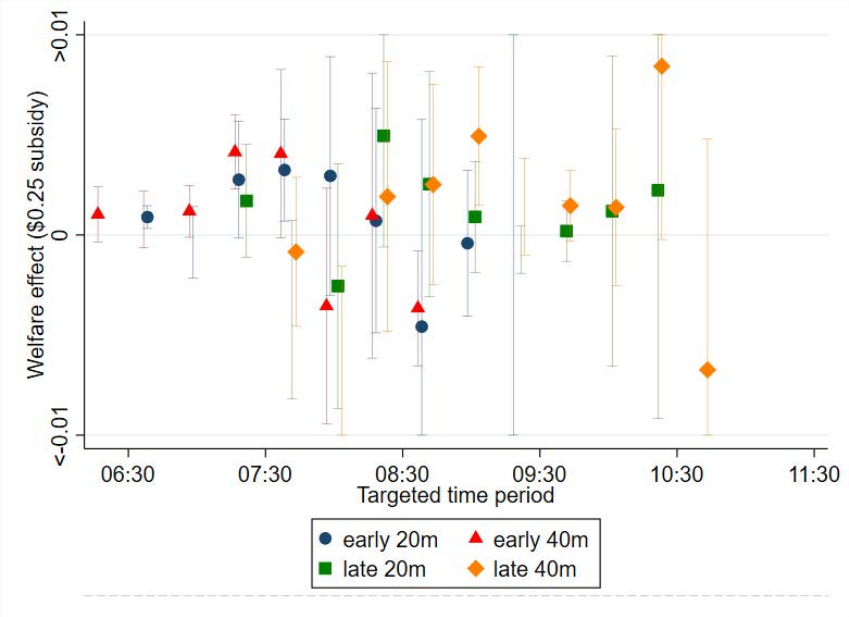
092000	084000	-0.02	0.03	1.04	1.72	0.00	0.45	0.49	0.72
092000	090000	-0.02	0.01	1.56	1.94	0.34	0.59	1.06	1.00
092000	094000	-0.02	0.01	0.93	0.28	-0.08	-0.52	1.11	1.01
092000	100000	-0.02	0.01	0.77	-0.16	-0.19	-0.81	1.32	1.12
094000	100000	-0.01	0.01	-0.02	-0.27	-0.71	-0.88	1.08	1.14
094000	102000	-0.01	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,

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

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 $V_j(p,s,e) = \max_{x_1^j, x_2^j} u(x_1^j, x_2^j, e) + Z - px_1^j - (p-s)x_2^j$ 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

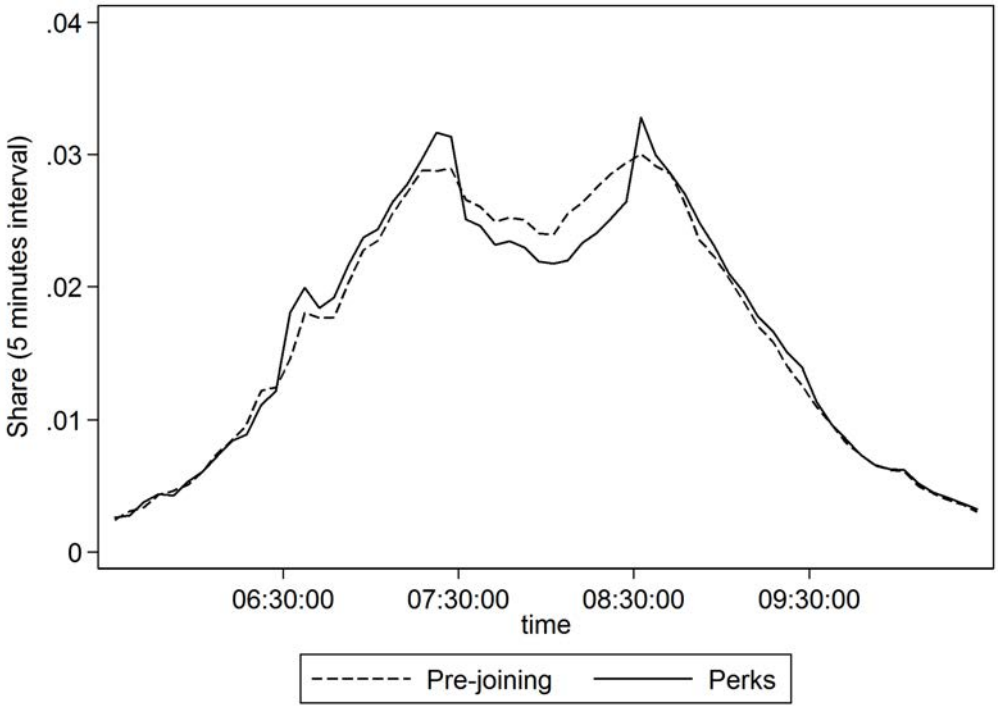
$$\begin{aligned}\frac{1}{N}(W(s') - W(0)) &= \frac{(n_a + n_b k)}{N}(p - c - MEC_1)\frac{dx_1^a}{ds}s' + \frac{(n_a + n_b k)}{N}(p - c - \frac{1}{2}s')\frac{dx_2^a}{ds}s' \\ &= \mu((p - c - MEC_1)\frac{dx_1^a}{ds}s' + (p - c - \frac{1}{2}s')\frac{dx_2^a}{ds}s') \quad \text{(A24)}\end{aligned}$$

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.

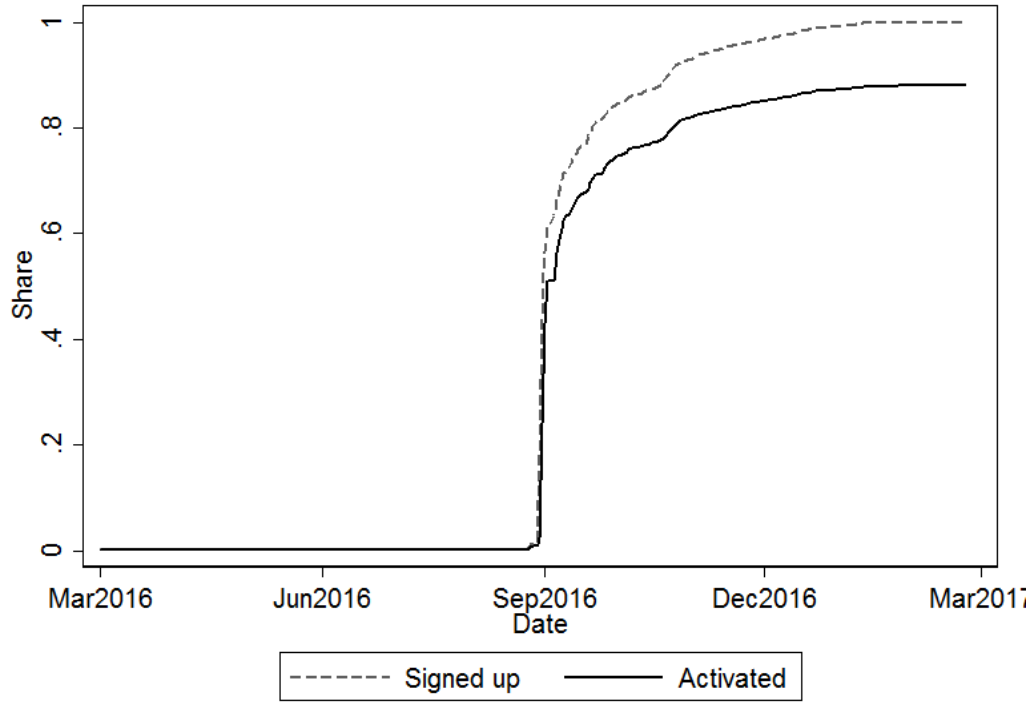
B Additional Figures and Tables

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



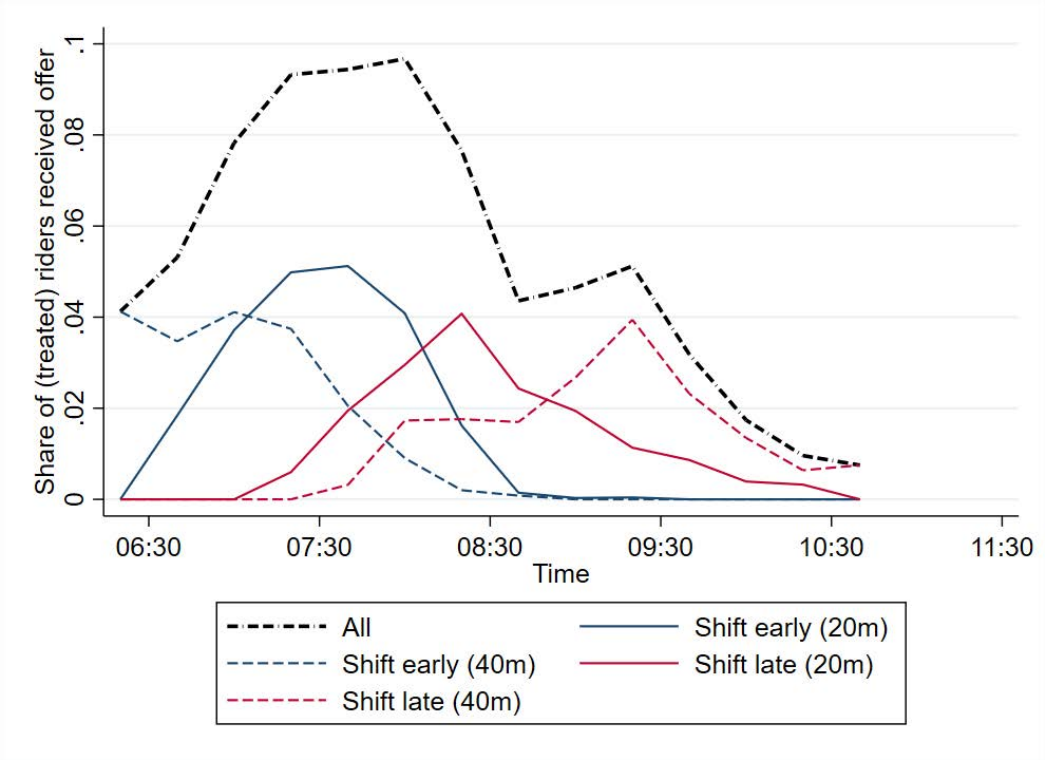
Note: The graph plots the cumulative share of users signed up, and who had activated.

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.

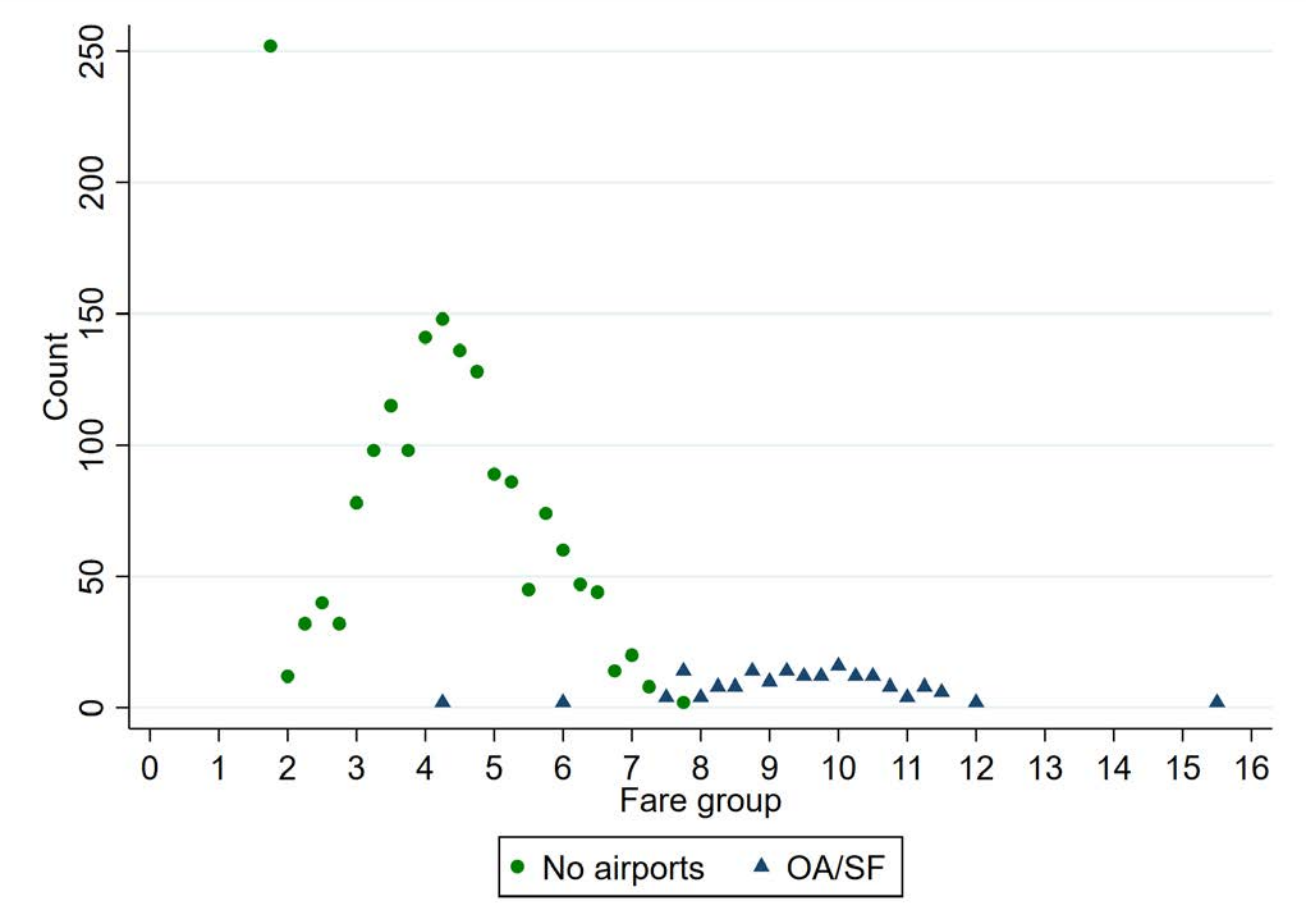
Table A13: Route level difference-in-differences estimate

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

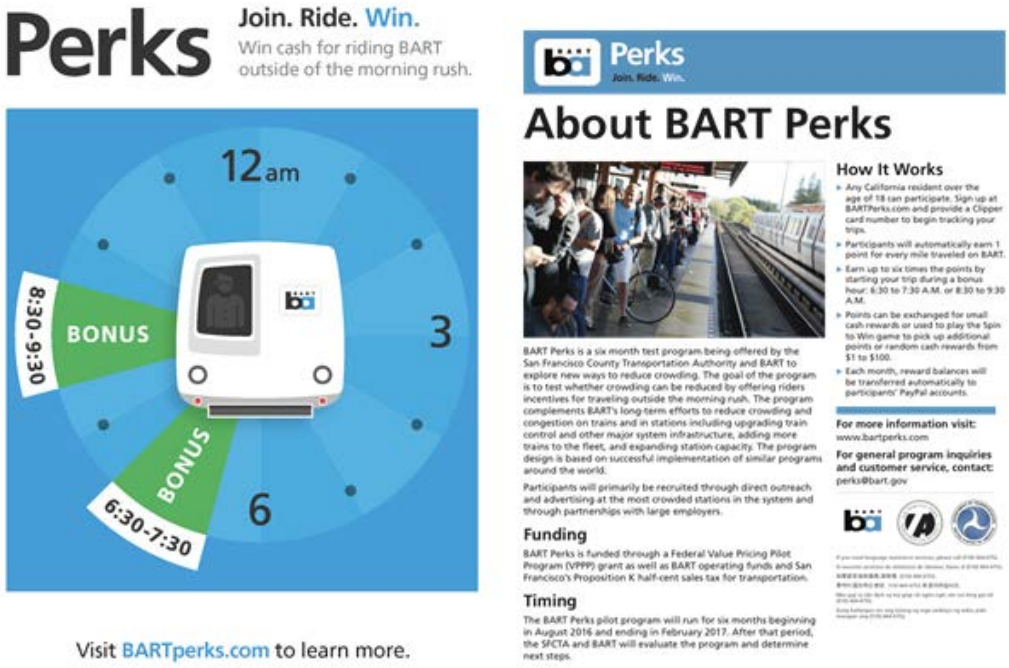
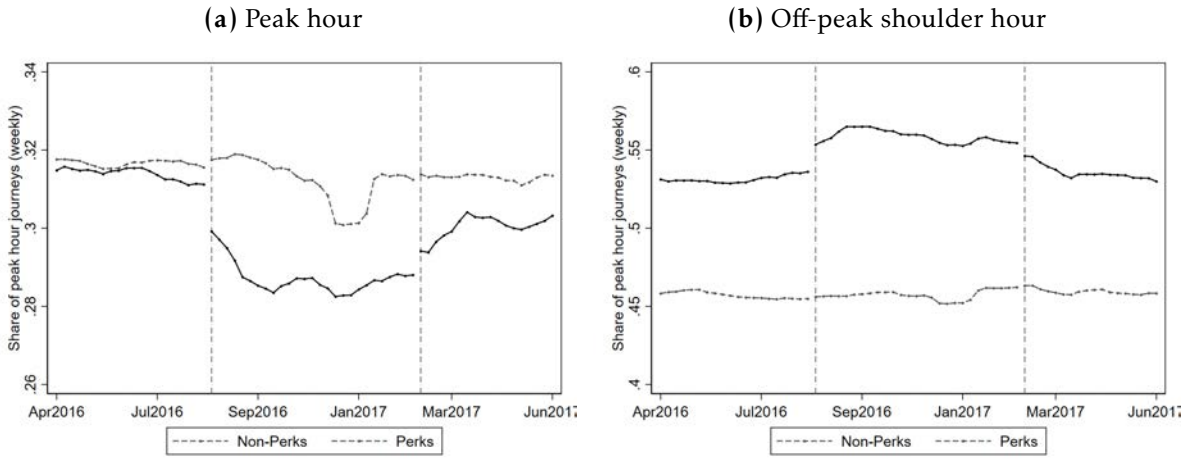


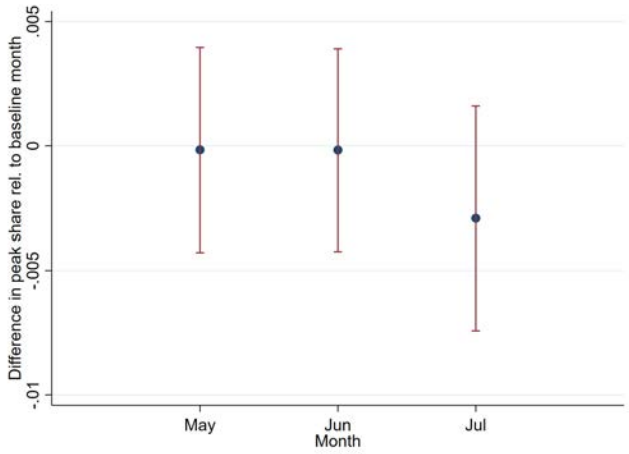
Figure A11: Long term - Share of peak hour journeys



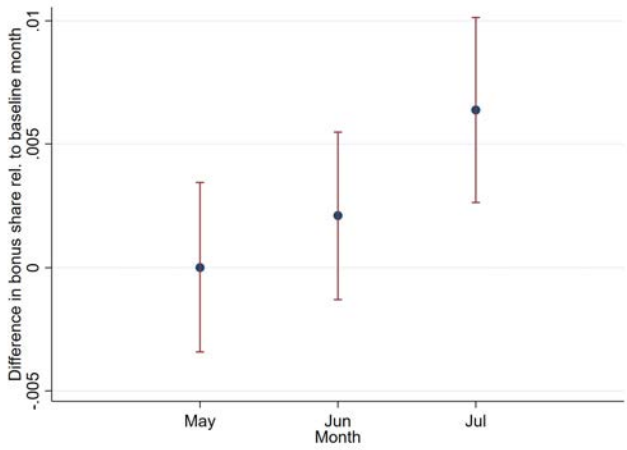
The graph plot trend of weekly average peak hour (panel (a)) and off-peak shoulder hour (panel (b)) journey share in the morning (6:30am-7:30am and 8:30am-9:30am among all journeys in 5:30am-10am), for weekdays. Sample include 1 Apr 2016-30 Jun 2017, excluding public holidays.

Figure A12: Test of pre-trend

(a) Share of peak hour



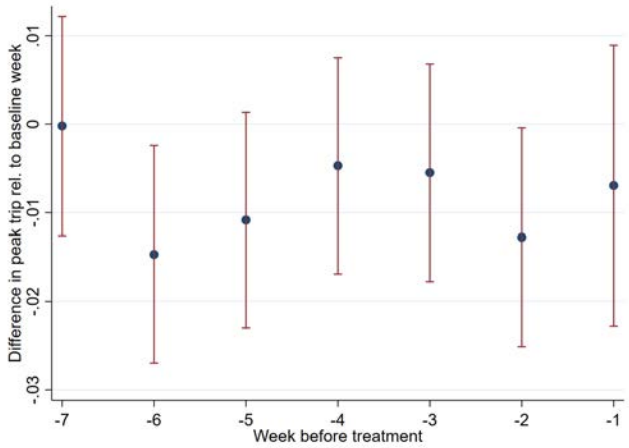
(b) Share of off-peak shoulder hour



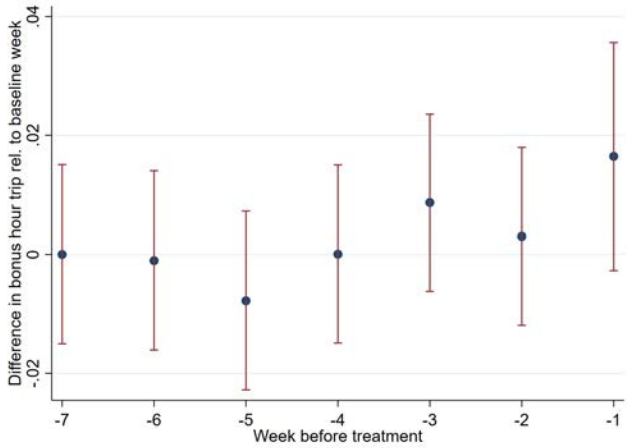
The figure plots the test of pre-trend with respect to Figure 1 and 2. It estimate a variant of equation 2, which is $y = \beta_1 * Treated * May + \beta_2 * Treated * June + \beta_3 * Treated * Jul + \eta_g + \gamma_t + \epsilon_{gt}$, where y_{gt} is the share of peak hour trips (panel (a)) or share 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

(a) Peak hour trips

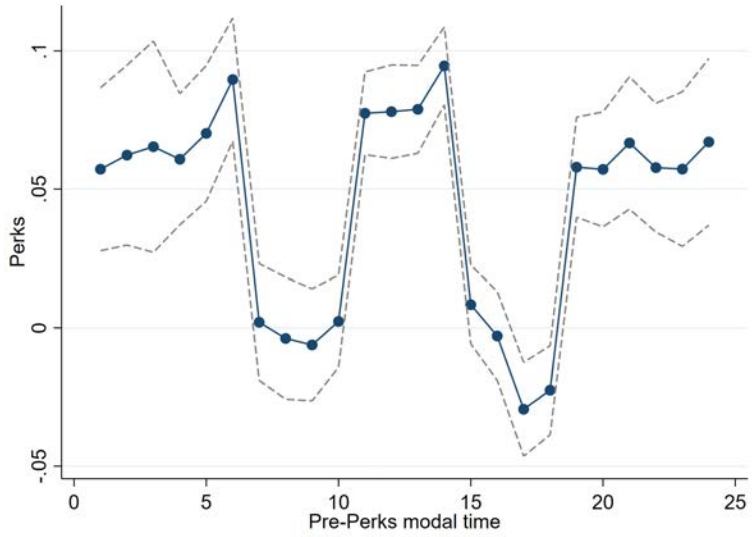


(b) Share of off-peak shoulder hour



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

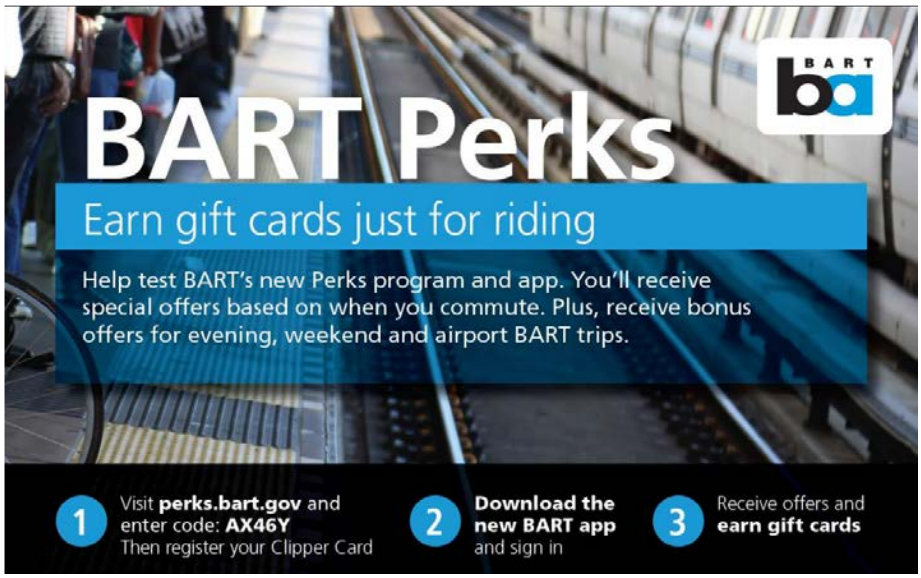
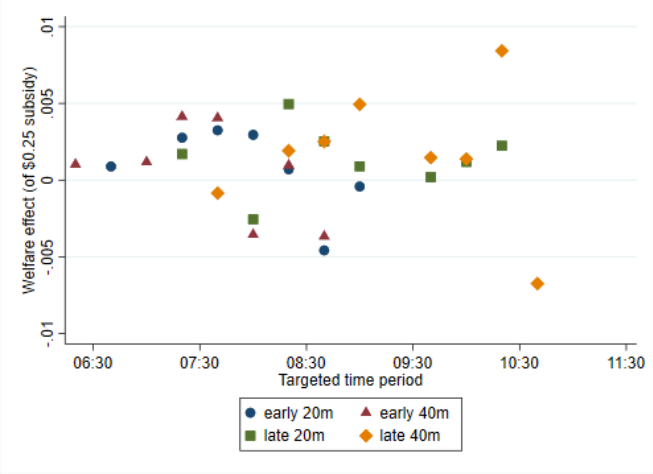
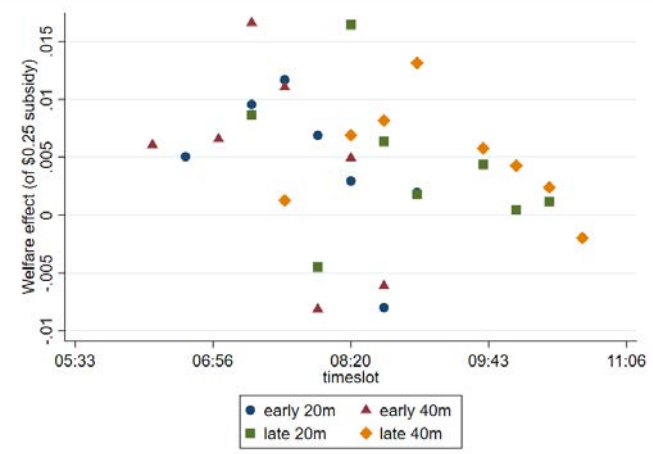


Figure A16: Welfare effect of subsidy in the field experiment – by subsidized time

(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

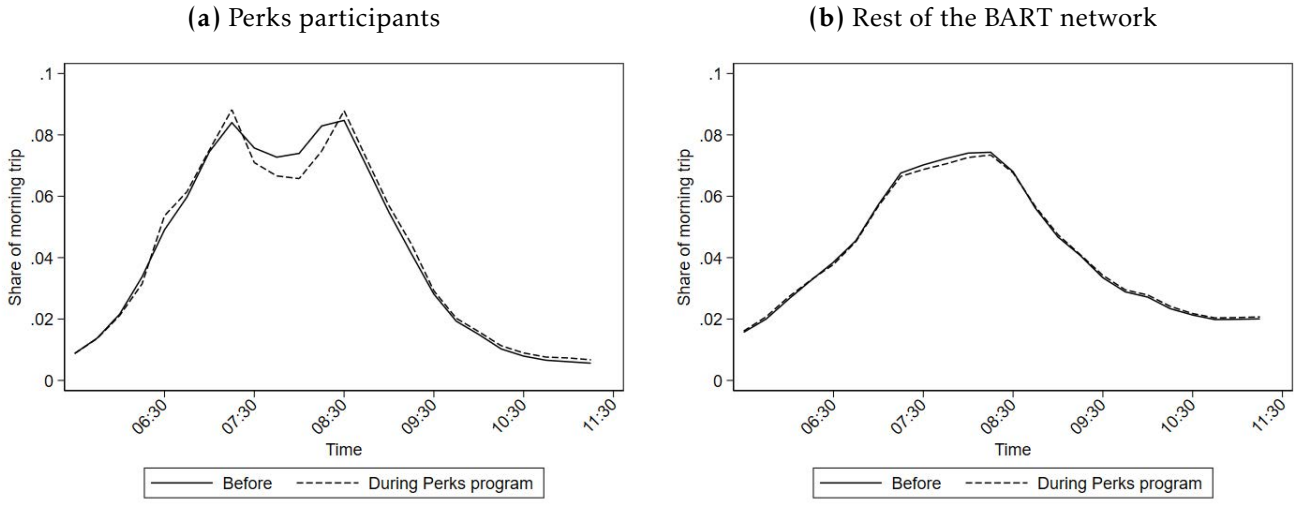
Notes: The table reports estimation of equation 2 at route level at weekly level. The outcome is weekly observation for each route in the BART network for Perks users or the rest of network. Sample includes 1 April 2016 to 28 Feb 2017, all weekdays excluding specific public holidays. Standard error clustered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



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

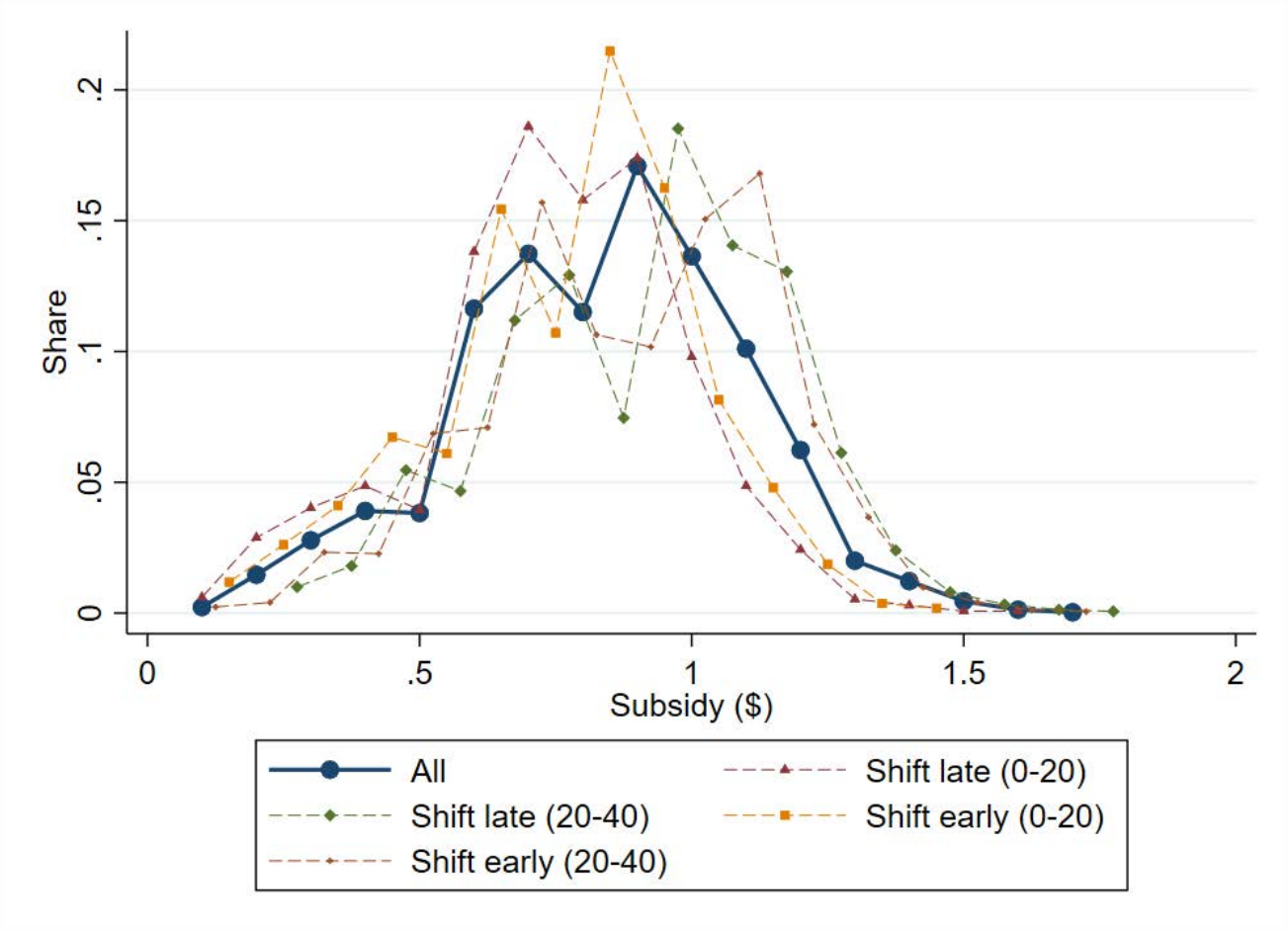
Notes: The table report summary statistics for experiment 2. Column (1) report the participants average daily number of trip by 20 minutes time interval in period before the experiment. Column (2) and (3) report on average for each date during the experiment period, the number of user received offer 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



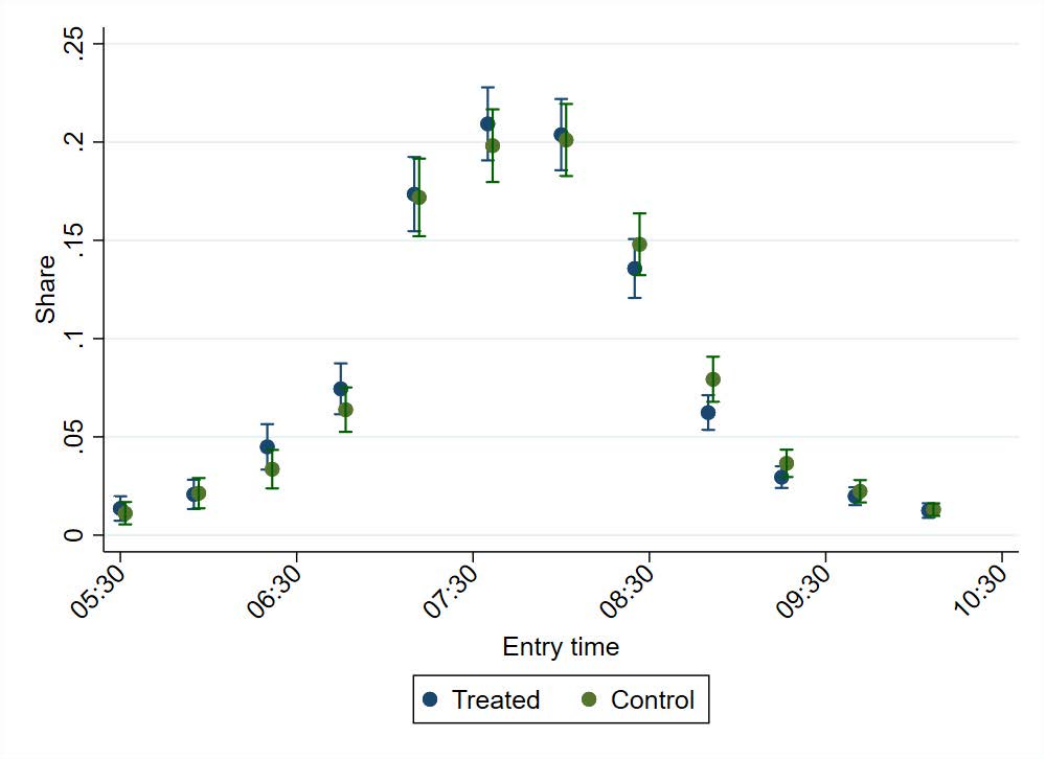
Notes: The figure plots the distribution of subsidy amount in experiment 2 for the morning subsidies. The y-axis represent the share of subsidy given with the amount (in bin width of 0.1$), among all subsidies in the respective type.

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.

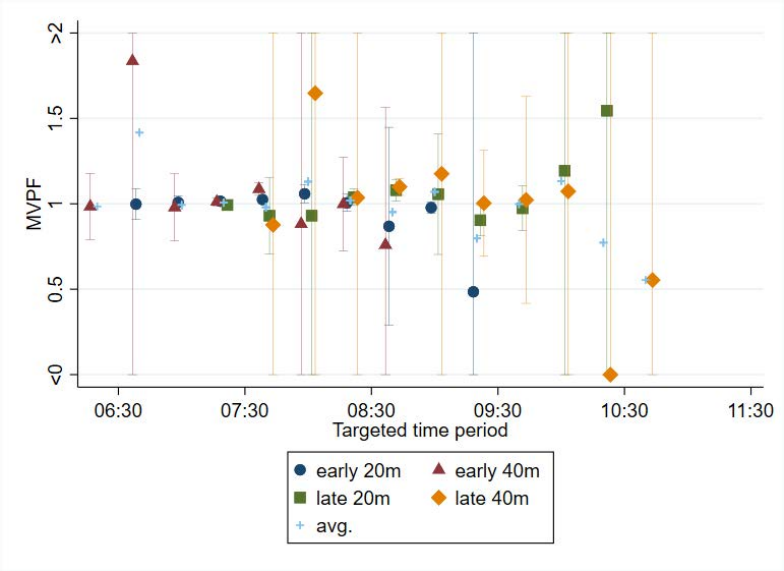
Table A22: Comparison of treatment and control group for experiment

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

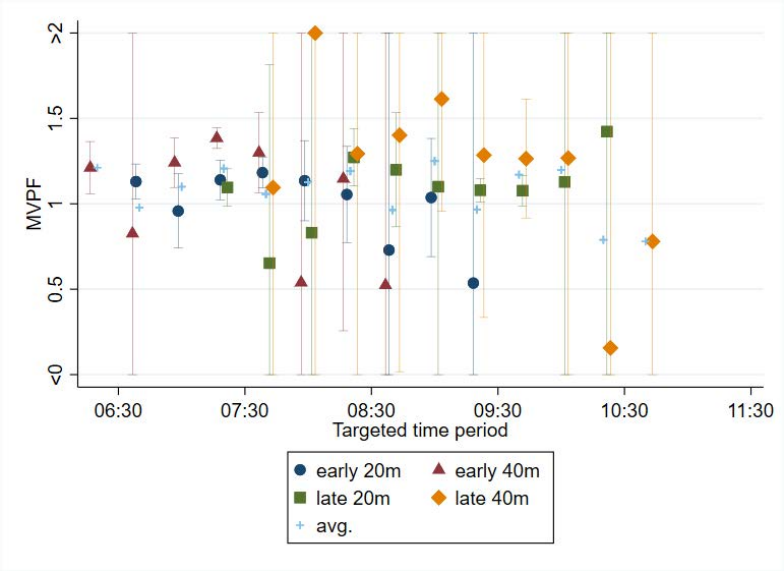
Notes: The table presents a regression of the outcomes on a treatment indicator for the field experiment, for the period three months before the experiment.

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.