TI 2024-081/VIII
Tinbergen Institute Discussion Paper

Using Nudging Information to Manage Congestion and Emissions in a Road and Metro Network

Zhiyuan Liang[1]
Vincent A.C. van den Berg[2]
Erik T. Verhoef[3]
Yacan Wang[4]

1 Beijing Jiaotong University

2 Vrije Universiteit Amsterdam, Tinbergen Institute

3 Vrije Universiteit Amsterdam, Tinbergen Institute

4 Beijing Jiaotong University

Using Nudging Information to Manage Congestion and Emissions in a Road and Metro Network

Zhiyuan Liang[a,b], Vincent A.C. van den Berg[b,c], Erik T. Verhoef[b,c], Yacan Wang[a*]

[a] *Department of Economics, School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China*

[b] *Department of Spatial Economics, SBE, VU Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*

[c] *Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands*

Abstract

This study studies nudging information as a strategy that can complement or substitute externality pricing, by influencing commuter behavior through awareness of the health and environmental impacts of their choices. We develop a bi-modal model with road and metro commuters, with bottleneck congestion on the road and dynamic crowding congestion in the metro. The model further incorporates health costs and environmental externalities, particularly for road commuters. When commuters are homogeneous, our findings indicate that nudging information generates positive welfare effects except in scenarios with extremely high crowding effects in the metro system. Moreover, nudging information can consistently complement flat road tolls by integrating information and toll schemes to enhance the system's social welfare impact. By adding heterogeneity in environmental preferences, car types, and income, the study further highlights that the effectiveness of such strategies depends on the varied behavioral responses from diverse individuals. Even when the crowding effect is relatively small with heterogeneity, nudging information may result in negative welfare effects by causing welfare-reducing swaps in road commuters' departure patterns; in such cases, it fails to complement flat tolls effectively.

Keywords: Congestion; Emissions; Nudging information; Bi-modal; Heterogeneity

1. Introduction

Urban areas worldwide are increasingly grappling with air pollution, which negatively impacts human health by causing respiratory and cardiovascular diseases, aggravating existing health conditions, and reducing life expectancy. Road traffic emissions have been identified as a significant contributor to urban air pollution, particularly in densely populated megacities with high vehicle concentrations (Vosough et al., 2022). In response, transportation authorities and policymakers have implemented various strategies to reduce traffic emissions and their associated externalities. One widely researched approach involves pricing schemes, such as congestion pricing and emission tolls, designed to encourage road users to reduce their travel during peak hours or switch to greener modes of transport, like public transit or cycling (de Palma and Lindsey, 2011). These schemes work by internalizing traffic emissions' external costs, making driving more expensive under certain conditions or in specific areas. The economic principle is straightforward: by increasing the cost of driving, these measures aim to reduce road demand and, consequently, overall emission levels.

While monetary incentives like road pricing have proven effective in altering travel behavior, they are with limitations. For instance, such measures can be perceived as punitive, disproportionately affecting lower-income individuals who may rely on private vehicles for commuting and lack viable alternatives (Van den Berg and Verhoef, 2011). They can also be hard to implement politically and expensive to operate.

As such, there is growing interest in exploring alternative approaches that can complement or replace traditional pricing mechanisms. Research has shown that nudging information can influence individuals' behavior effectively, particularly when encouraging sustainable practices (Sudarshan, 2017; Myers & Souza, 2020). Unlike temporary monetary incentives, under multiple rounds of intervention, nudging can instill long-lasting habits by fostering a deeper understanding of the consequences of one's actions (Guerassimoff and Thomas, 2015; Marchiori et al., 2017). For example, informing drivers about the exact amounts of emissions their vehicle produces per trip can motivate them to consider alternative modes of transportation, especially if they are environmentally conscious.

However, the effectiveness of nudging strategies may not be uniform across all individuals. Studies have found that people's responses to nudging information vary based on personal values, environmental awareness, and social norms (Erlei, 2008; Kamas & Preston, 2012; Lucas et al., 2014). While some individuals may be highly motivated by environmental concerns and readily alter their behavior in response to nudging, others may be less affected or even indifferent to such information. This variation in responses highlights the need to consider the diversity of commuters' environmental preferences when designing and implementing nudging interventions.

Against this background, this study examines the efficiency and distributional impacts of nudging information in a bi-modal setting with dynamic congestion, allowing for heterogeneity in commuters' environmental preferences. This heterogeneity can include cases where some drivers completely ignore the information. We address the following research questions:

1) How does nudging information affect individuals' mode and departure time choices with homogeneous and heterogeneous commuters?
2) What is the relationship between commuters' heterogeneity in environmental preferences, income, car type, and the distributional impacts of nudging information?
3) Is nudging information always a welfare-improving strategy when considering different types of commuters' heterogeneities? Can it be considered a complement to tolling schemes?

To accomplish this, we establish a bi-mode traffic model, including road and metro commuters, by incorporating the road bottleneck model proposed by Vickrey (1969) and the dynamic metro model proposed by de Palma et al. (2017). Various two-mode traffic models are proposed in the literature for investigating travelers' behavior and regulators' optimal pricing strategy for internalizing the externalities during traveling (e.g., Mirabel and Reymond, 2011; Li et al., 2012; Van den Berg and Verhoef, 2014; Li and Zhang, 2020). Different from studies focused solely on the auto-only system, the interaction and substitution effects between different modes (notably auto and metro) directly and indirectly affect the cost of travel by a particular mode and thus, the mode choices of travelers in the two-mode travel system, which should not be ignored when designing an optimal pricing scheme (Li and Zhang, 2020). Regarding methodology, static and dynamic models of congestion have been developed. Compared to static congestion model work, the dynamic congestion model (e.g., the classic bottleneck model) has an advantage in investigating commuters' departure time choices. In light of the above discussion, this paper develops a two-mode model considering dynamic congestion in both modes (auto and metro). We can thus consider not only mode choice but also measure both the departure time choice for both modes, under queuing for cars and in-vehicle passenger crowding externality in the metro system, and account for all these factors in welfare assessment.

A large body of literature on road congestion also suggests the need to consider emission externalities during travel. Most studies focus on optimizing emission pricing strategies, identifying these as effective solutions for addressing environmental problems caused by car emissions in most cases. According to the modeling approaches in these studies, different emission cost functions for road commuters have been adopted. For a general static road model, emission cost is usually assumed to vary with the distance and the travel time through the link (Ma et al., 2017). By further treating commuters' departure time endogenously, drivers' emission cost has also been related to the queuing time at the bottleneck in earlier dynamic congestion models (e.g., Liu et al., 2015; Xiao et al., 2016; Coria and Zhang, 2017). Furthermore, a series of studies consider more complex emission cost functions. In these analyses, the vehicle's speed, wind speed, wind direction and more can be included in the emission cost function. The pricing of these emissions has also been studied (e.g., Vosough et al., 2022). We will assume that the emission rate varies under different driving conditions during travel, distinguishing cruising on the freeway versus queuing at the bottleneck. The metro system is assumed to be a relatively greener travel mode, which we assume does not generate any emission costs in this paper. This assumption is commonly used when comparing road and public transport travel modes (e.g., Carroll et al., 2019; Thomas and Serrenho, 2024).

Previous studies have considered information schemes aimed at reducing the welfare losses caused by travel time uncertainty (e.g., Emmerink et al., 1996; Verhoef et al., 1996a; de Palma and Lindsey, 1998; Liu et al.,2020; Yu et al., 2021; Yu et al., 2023). Unlike information about varying conditions on the roads caused by, for example, traffic accidents in the above studies, this paper concerns the emission and health costs generated during driving. We assume these are initially not fully perceived by drivers. The associated informational incentives are called nudging information in the behavioral economics literature (Bhargava and Loewenstein, 2015). Based on behavioral theory, we allow nudging information to change individuals' behavior. We assume this happens systematically as the nudging information fills an initial knowledge deficit. Our extended model also allows for heterogeneity in responses to nudging and in preferences. To the best of our knowledge, we are the first to explore the impact of nudging information on the internalization of emissions and congestion in traffic modeling.

Our primary methodological contribution lies in studying nudging information aimed at internalizing the emission externalities and analyzing it in a dynamic user equilibrium model with road and metro travel with congestion, crowding and environmental externalities. Our approach considers the heterogeneity of commuters and imperfect substitutability between car and metro travel modes. These are complicating factors that are, however, crucial to consider in establishing the welfare impacts of nudging information. By capturing these intricacies, our model can investigate the complementary and substitutive relationships between nudging information and tolling.

The remainder of the paper is organized as follows. Section 2 introduces the benchmark model, and Section 3 analyzes its results. Section 4 adds heterogeneity. Section 5 presents a numerical model that illustrates the effects. As there is substantial uncertainty in empirical values for the parameters, Section 6 does extensive sensitivity checks. Finally, Section 7 concludes.

2. Benchmark model

2.1 Setting the stage

Consider a continuous flow of commuters traveling from home to work each morning. For now, assume that everyone has identical preferences on the timing and cost of their trips. Two modes can be used: a road and a metro link, where the road experiences bottleneck congestion, while the metro system encounters dynamic crowding. The choice between modes depends on the travel costs of each option. Let N_A and N_M represent the number of auto and metro commuters, respectively.

Solo-driving auto commuters travel through a bottleneck at the end of the road, with a deterministic capacity of s (passenger cars per unit of time). A queue develops when the arrival rate of auto commuters exceeds the road capacity. For conventional linear schedule delay costs, the resulting cost function of auto commuters is shown below:

$$C_A(t_d) = \alpha\left(T_{Af} + q_A(t_d)\right) + \begin{cases} \beta\left(t^* - T_{Af} - t_d - q_A(t_d)\right) & t^* \geq T_{Af} + t_d + q_A(t_d) \\ \gamma\left(T_{Af} + t_d + q_A(t_d) - t^*\right) & t^* \leq T_{Af} + t_d + q_A(t_d) \end{cases} \tag{1}$$

where $C_A(t_d)$ and $q_A(t_d)$ denotes the auto commuters' travel cost and the travel time through the bottleneck who depart from home at time t_d, respectively; T_{Af} denotes the fixed in-vehicle time from home to the bottleneck; t^* denotes the official starting time, α denotes the value of travel time; and β and γ denotes the values of schedule delay early and late, respectively.

Based on the existing literature on road bottlenecks (e.g., Arnott et al., 1990), it is straightforward to derive the generalized travel cost for auto commuters without any incentive policies, as follows:

$$C_A = \alpha T_{Af} + \frac{\beta\gamma}{\beta+\gamma}\frac{N_A}{s} \tag{2}$$

Metro commuters travel from a subway station; for simplicity, the travel time from home to the subway station is disregarded. There are M service runs in total, departing from the subway station each h minutes and with a capacity of c. The travel cost function of metro commuters is:

$$C_M(t_k) = \alpha T_{Mf} + g(n_k) + \sigma^k \tag{3}$$

$$\sigma^k = \begin{cases} \beta(t^* - T_{Mf} - t_k - q_M(t_k)) & t^* \geq T_{Mf} + t_k + q_M(t_k) \\ \gamma(T_{Mf} + t_k + q_M(t_k) - t^*) & t^* \leq T_{Mf} + t_k + q_M(t_k) \end{cases} \tag{4}$$

where $C_M(t_k)$ and σ^k denotes the metro commuters' travel cost and schedule delay cost who take the service run k departs at time t_k, respectively, T_{Mf} denotes the fixed in-vehicle time in metro mode, and $g(n_k)$ denotes crowding or body congestion cost. The latter is a monotonically increasing function of the number of passengers n_k in service run k.

The body congestion cost function $g(n_k)$ is assumed to be linear:

$$g(n_k) = g\frac{n_k}{c} \tag{5}$$

Hence, the equilibrium generalized travel cost for a metro commuter can be obtained:

$$C_M = \sigma + \alpha T_{Mf} + \frac{gN_M}{Mc} \tag{6}$$

$$\sigma = \frac{1}{M}\sum_{k=1}^{M}\sigma^k \tag{7}$$

where σ denotes the unweighted average schedule delay cost for metro users. [1]

---

[1] As previously noted, this analytical model focuses on road and metro systems. However, the metro component can be extended to represent any transit mode that operates independently of road traffic, such as fully separated BRT systems, monorails, or urban trains, but not buses or trams when being part of the congested road system.

2.2 The emission cost of private cars

As the emission rate varies over driving conditions (Liu et al., 2015), the car emissions in this paper are categorized into two aggregate conditions: cruising on the freeway (at a rate λ_f per unit of time) and queuing at the bottleneck (at a rate λ_q per unit of time). Given that emissions in a queuing state are higher than in a free-flow state, it is assumed that $\lambda_q > \lambda_f$.

The emission of a car departing at time t_d is:

$$e_A(t_d) = \lambda_f T_{Af} + \lambda_q q_A(t_d) \tag{8}$$

We assume this emission cost of car use is a pure externality and thus would be fully ignored by drivers when they do not receive any further information during traveling.

2.3 The health cost of occupying vehicles

Prolonged time spent in cars can lead to obesity (Frank et al., 2004; Sugiyama et al., 2016) and mental health issues (Wang et al., 2019). We therefore assume that there is a health cost associated with being stuck in a car, which we assume is linearly related to both free-flow travel time and queuing time. This relationship can be expressed as follows:

$$H_A(t_d) = \lambda_h (T_{Af} + q_A(t_d)) \tag{9}$$

where λ_h denotes the constant shadow cost of health related to the travel time by car.

In general, drivers may already regard part of this health cost as a relevant component of travel price without information incentives. We assume that this concerns a fraction $\mu_h (0 \leq \mu_h \leq 1)$. The perceived health price during "uninformed" driving can be then expressed as follows:

$$H_{AP}(t_d) = \mu_h H_A(t_d) \tag{10}$$

The departure rate of auto commuters and the start and end of departure time without information incentives can then be derived as:

$$r(t_d) = \begin{cases} s + \frac{\beta s}{\alpha - \beta + \mu_h \lambda_h} & t^* \geq T_{Mf} + t_k + q_M(t_k) \\ s - \frac{\gamma s}{\alpha + \gamma + \mu_h \lambda_h} & t^* \leq T_{Mf} + t_k + q_M(t_k) \end{cases} \tag{11}$$

$$\begin{cases} t_s = t^* - T_{Af} - \frac{\gamma}{\beta + \gamma} \frac{N_A}{s} \\ t_e = t^* - T_{Af} + \frac{\beta}{\beta + \gamma} \frac{N_A}{s} \end{cases} \tag{12}$$

Hence, the generalized travel price in the road bottleneck without information incentives can be derived as follow:

$$P_A = (\alpha + \mu_h \lambda_h) T_{Af} + \frac{\beta \gamma}{\beta + \gamma} \frac{N_A}{s} \tag{13}$$

2.4 Nudging information for auto commuters

We assume that the regulator can provide two types of information to encourage drivers to internalize emission costs.

The first concerns the actual health cost when being stuck in the vehicle. We assume that all travelers care about their health, and we assume that under information incentives, drivers consider the full, actual health costs communicated to them, and treat these entirely as private costs. We thus ignore that part of health costs that people may ignore in the face of health insurance, and we also ignore the possibility that information may be only partly successful.

The second type of information pertains to the emission cost generated during the car trip. In contrast, when informed about emissions, drivers might recognize this latent cost $e_A(t_d)$, and incorporate it only partly into their perceived travel cost, as follows:

$$e_{AP}(t_d) = \mu_e e_A(t_d). \tag{14}$$

Here, $e_{AP}(t_d)$ denotes the emission price perceived and accounted for by drivers. , and μ_e represents the fraction that is considered, which is a parameter related to the individual's environmental preferences ($0 \leq \mu_e \leq 1$).

With nudging information, road commuters thus perceive the full actual health and part of the emission costs of their travel. As a result, they choose their departure time while considering not only their travel time and schedule delays but also the health and internalized emission costs. The travel cost for an auto commuter departing at time t_d

can then be represented as follows:

$$C_A(t_d) = (\alpha + \lambda_h)\left(T_{Af} + q_A(t_d)\right) + \mu_e e_A(t_d) + \begin{cases} \beta(t^* - T_{Af} - t_d - q_A(t_d)) & t^* \geq T_{Af} + t_d + q_A(t_d) \\ \gamma(T_{Af} + t_d + q_A(t_d) - t^*) & t^* \leq T_{Af} + t_d + q_A(t_d) \end{cases} \tag{15}$$

Similar to the equilibrium derived from the basic bottleneck model without nudging information, the departure rate of auto commuters, as well as the start and end times of departure in user equilibrium, can now be derived as follows:

$$r(t_d) = \begin{cases} s + \frac{\beta s}{\alpha - \beta + \lambda_h + \mu_e \lambda_q} & t^* \geq T_{Af} + t_d + q_A(t_d) \\ s - \frac{\gamma s}{\alpha + \gamma + \lambda_h + \mu_e \lambda_q} & t^* \leq T_{Af} + t_d + q_A(t_d) \end{cases} \tag{16}$$

$$\begin{cases} t_s = t^* - T_{Af} - \frac{\gamma}{\beta+\gamma}\frac{N_A}{s} \\ t_e = t^* - T_{Af} + \frac{\beta}{\beta+\gamma}\frac{N_A}{s} \end{cases} \tag{17}$$

The generalized travel price on the road under nudging information can now be shown to be:

$$P_A^I = (\alpha + \lambda_h + \mu_e \lambda_f) T_{Af} + \frac{\beta\gamma}{\beta+\gamma}\frac{N_A}{s} \tag{18}$$

The cost components covered by information vary in proportion with travel delays. In the basic bottleneck model, α does not affect peak duration nor peak start and ending times. Similarly, in our model, the introduction of the nudging information also leaves these unaltered. The intuition is that queuing delays through departure time adjustments will respond inversely proportional to changes in α.

2.5 Demand side

Finally, we consider the price sensitivity of commuters, for both auto and metro, by applying a quadratic total travel benefit function:

$$U(N_A, N_M) = a_A N_A + \frac{1}{2} b_A {N_A}^2 + a_M N_M + \frac{1}{2} b_M {N_M}^2 + \eta N_A N_M \tag{19}$$

This functional form leads to the following linear inverse demand functions as follows:

$$D_A = a_A + b_A N_A + \eta N_M \tag{20}$$

$$D_M = a_M + b_M N_M + \eta N_A \tag{21}$$

where D_A and D_M represent the inverse demand functions for auto and metro users, respectively, thus indicating the marginal willingness to pay for the $N_A th$ auto user and the $N_M th$ metro user. The parameters a_A and a_M are the intercepts, representing maximum willingness-to-pay. For both modes, the coefficients η measure by how much the inverse demand decreases as the number of users on the other mode increases and thus captures mode substitution. For metro users, the coefficients b_M measure the extent to which D_M decreases as more users choose the metro, and similarly, for b_A for auto users.

3. Basic equilibria under different incentives

We are now ready to characterize equilibrium under different incentive settings.

3.1 No incentives

First, we discuss user equilibrium without any intervention to metro and auto mode commuters. Hence, the user mode choice equilibrium condition can be written as:

$$P_A = D_A(N_A, N_M) \tag{22}$$

$$C_M = D_M(N_M, N_A) \tag{23}$$

The number of auto and metro commuters in user equilibrium N_A^N and N_M^N can then be derived analytically (see Appendix A). Note that equation (23) shows that we assume average cost pricing for metro which, give our assumptions, implies a zero fare for public transport. A simultaneous introduction of a fare equal to average supplier cost, and an equivalent upward shift of a_M would leave the analysis unaffected.

3.2 Nudging information

Next, we discuss the equilibrium under nudging information to drivers. The equilibrium condition can be written as:

$$P_A^I = D_A(N_A, N_M) \tag{24}$$

$$C_M = D_M(N_M, N_A) \tag{25}$$

The number of auto and metro commuters under nudging information N_A^I and N_M^I can again be derived analytically (see Appendix A).

3.3 Second-best flat road toll

As a useful second-best benchmark, we now consider the case where the government can implement uniform external toll for auto commuters, to maximize in a second-best fashion social welfare in this bi-modal system. We use social surplus as our measure for social welfare, which can be written as:

$$SW = \int_{(0,0)}^{(N_A,N_M)} U(N_A, N_M) - TP_A - TC_M - TE_A - (1-\mu_h)TH_A \tag{26}$$

where TC_A ,TE_A and TH_A denotes the total travel cost, total emission cost and total health cost for drivers, respectively.

The equilibrium conditions for optimal road toll can be written as:

$$P_A^T = P_A + \tau_A = D_A(N_A, N_M) \tag{27}$$

$$C_M = D_M(N_M, N_A) \tag{28}$$

Hence, the number of auto and metro commuters under second-best optimal uniform road pricing ${N_A}^O$ and ${N_M}^O$ can be derived based on the above equilibrium conditions (see Appendix A).

The second-best flat road toll can then be derived as:

$${\tau_A}^O = N_A \frac{\partial P_A}{\partial N_A} + \frac{\partial TE_A}{\partial N_A} + (1-\mu_h)\frac{\partial TH_A}{\partial N_A} - \frac{\partial D_M}{\partial N_A} * \left(\frac{D_M - C_M - \frac{\partial C_M}{\partial N_M} N_M}{\frac{\partial D_M}{\partial N_M} - \frac{\partial C_M}{\partial N_M}}\right) \tag{29}$$

The second-best flat road toll includes, besides the conventional marginal external cost of congestion and emissions, a term reflecting the otherwise ignored health cost and a final term that will be negative when the metro is priced below marginal social cost. The weight is a generalization of the term appearing in the second-best toll for the classic two-route problem (Verhoef et al., 1996b), where the weight reflects the substitutability between modes; see also Small et al. (2024, Section 4.5.1).

3.4 Second-best flat road toll with nudging information

Next, we consider a mixed second-best policy design aimed at maximizing social welfare, where the government can implement a uniform external road toll while also providing nudging information to each auto commuter. We use social surplus as the measure of social welfare, which can be represented as follows:

$$SW = \int_{(0,0)}^{(N_A,N_M)} U(N_A, N_M) - TP_A^I - TC_M - (1 - \mu_e)TE_A \tag{30}$$

The equilibrium conditions become:

$$P_A^{IT} = P_A^I + \tau_A = D_A(N_A, N_M) \tag{31}$$

$$C_M = D_M(N_M, N_A) \tag{32}$$

Based on the equilibrium conditions, we can determine the number of commuters in each mode ${N_A}^*$ and ${N_M}^*$ in equilibrium (see Appendix A).

With nudging information, the second-best flat road toll can be written as:

$${\tau_A}^* = N_A \frac{\partial P_A^I}{\partial N_A} + (1 - \mu_e) \frac{\partial TE_A}{\partial N_A} - \frac{\partial D_M}{\partial N_A} * \left(\frac{D_M - C_M - \frac{\partial C_M}{\partial N_M} N_M}{\frac{\partial D_M}{\partial N_M} - \frac{\partial C_M}{\partial N_M}}\right) \tag{33}$$

The interpretation is similar to that of the toll in eq. (30), with the key differences being that no longer a toll component for private health cost is needed; and, a smaller toll component is needed for the environmental externality due to partial “self-internalization”.

Proposition 1. *When commuters are homogeneous, nudging information can generate positive welfare effects provided crowding externalities in public transport are not excessive.*

Proof. The proof of Proposition 1 is in Appendix B.

Proposition 1 highlights that while nudging information can encourage drivers to internalize health and emission costs associated with driving, it does not inherently guarantee a welfare improvement. The effect of nudging in part depends on the balance between the reduced congestion and emissions on the road and the potential increase in crowding within the metro system. As nudging shifts a larger portion of commuters from roads to metro, a trade-off emerges: while road congestion and emissions decrease, the metro experiences increased crowding costs. Therefore, the welfare outcome of nudging in part relies on the relative magnitude of these externalities in both modes, indicating that nudging information is most effective in scenarios where the metro's crowding externalities are smaller.

We can derive a sufficient but not necessary condition under which the nudging information always generates positive welfare effects:

Proposition 2. *In situations where a naive Pigouvian flat road toll consistently generates positive welfare effects, nudging information will always enhance overall social surplus under homogeneous preferences.*

Proof. The proof of Proposition 2 is in Appendix C.

We define the naive Pigouvian flat road toll for drivers can be written as:

$$\tau_A^{np} = N_A \frac{\partial P_A}{\partial N_A} + \frac{\partial TE_A}{\partial N_A} + (1-\mu_h)\frac{\partial TH_A}{\partial N_A} \tag{34}$$

It is naive in the sense that it ignores crowding congestion in the metro system and only seeks to price uninternalized externalities on the road. Proposition 2 implies that when such a naive Pigouvian flat road toll improves welfare, it also becomes inherently welfare-enhancing for the government to encourage a shift of drivers to the metro system through nudging information.

Proposition 3. *Nudging information will be complementary to the second-best flat road toll: the two instruments can jointly reach a higher social welfare level than the second-best flat alone when commuters are homogeneous.*

Proof. See Appendix D.

Proposition 3 suggests that nudging information complements the second-best flat road toll, enabling the system to attain a higher social welfare level than would be possible through the flat toll alone. This result is consistent with earlier findings on comparing between information and pricing in the literature, such as Verhoef et al. (1996b) who demonstrate that providing targeted information can complement second-best pricing strategies, under travel time uncertainty.

4. Heterogenous commuters

4.1 Heterogenous environmental preference

In reality, not all commuters are equally concerned about the emission costs from driving. Notably when commuters are heterogeneous in their environmental preferences, there can be varying responses to the emission information provided by the government.

To capture the impacts of such heterogeneity in a basic way, we will assume the existence of two groups of commuters, with high and low environmental preferences, respectively. The equilibrium numbers of car commuters in these two groups will be denoted as N_{AH} and $N_{AL}(N_{AH} + N_{AL} = N_A)$. Given the information provided, the travel prices for these two groups can be represented as follows:

$$P_{AH}(t_d) = \mu_{eH} e_A(t_d) + (\alpha + \lambda_h)\left(T_{Af} + q_A(t_d)\right) + \begin{cases} \beta(t^* - T_{Af} - t_d - q_A(t_d)) & t^* \geq T_{Af} + t_d + q_A(t_d) \\ \gamma(T_{Af} + t_d + q_A(t_d) - t^*) & t^* \leq T_{Af} + t_d + q_A(t_d) \end{cases} \tag{35}$$

$$P_{AL}(t_d) = \mu_{eL} e_A(t_d) + (\alpha + \lambda_h)\left(T_{Af} + q_A(t_d)\right) + \begin{cases} \beta(t^* - T_{Af} - t_d - q_A(t_d)) & t^* \geq T_{Af} + t_d + q_A(t_d) \\ \gamma(T_{Af} + t_d + q_A(t_d) - t^*) & t^* \leq T_{Af} + t_d + q_A(t_d) \end{cases} \tag{36}$$

where μ_{eH} and μ_{eL} denote the information perception parameters for commuters with high and low environmental preferences, respectively ($0 \leq \mu_{eL} \leq \mu_{eH} \leq 1$). For transparency of exposition, we for now assume all other time preference are equal between the two groups.

The generalized travel price on the road under information can now be derived:

$$P_{AL}^I = (\alpha + \lambda_h + \mu_{eL}\lambda_f)T_{Af} + \frac{\beta\gamma}{\beta+\gamma}\frac{N_{AL} + \frac{\alpha + \mu_{eL}\lambda_q + \lambda_h}{\alpha + \mu_{eH}\lambda_q + \lambda_h} N_{AH}}{s} \tag{37}$$

$$P_{AH}^I = (\alpha + \lambda_h + \mu_{eH}\lambda_f)T_{Af} + \frac{\beta\gamma}{\beta+\gamma}\frac{N_{AL} + N_{AH}}{s} \tag{38}$$

The user equilibrium for the two groups in the road bottleneck is illustrated in

Figure 1, where t_a temporally denotes the road commuters' arrival time. Due to the difference in the environmental preference parameter μ_e between the two groups, auto commuters are now separated under information incentives. This result mimics insights from previous bottleneck studies with heterogeneity in α (e.g., Van den Berg and Verhoef, 2011).

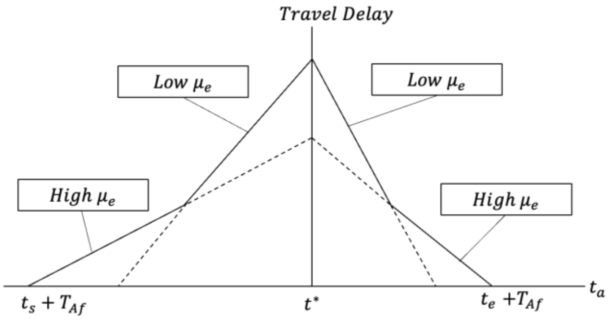


Figure 1. Travel delay for auto commuters with heterogenous environmental preferences under nudging information.

Similar to the homogeneous case, the equilibrium travel price without any incentives is:

$$P_{AL} = (\alpha + \mu_h \lambda_h) T_{Af} + \frac{\beta\gamma}{\beta+\gamma} \frac{N_{AL}+N_{AH}}{s} \tag{39}$$

$$P_{AH} = (\alpha + \mu_h \lambda_h) T_{Af} + \frac{\beta\gamma}{\beta+\gamma} \frac{N_{AL}+N_{AH}}{s} \tag{40}$$

With two groups of users, we apply the following four linear demands:

$$D_{AL} = a_{AL} + b_{AL} N_{AL} + \eta_L N_{ML} \tag{41}$$

$$D_{AH} = a_{AH} + b_{AH} N_{AH} + \eta_H N_{MH} \tag{42}$$

$$D_{ML} = a_{ML} + b_{ML} N_{ML} + \eta_L N_{AL} \tag{43}$$

$$D_{MH} = a_{MH} + b_{MH} N_{MH} + \eta_H N_{AH} \tag{44}$$

4.2 Heterogenous environmental preferences and car types

In this subsection, we further explore the links between commuters' environmental preferences and the types of cars. Similar to section 3.2, two discrete commuter groups with high and low environmental preferences are considered, while the equilibrium numbers of commuters in these two groups are denoted by N_{AH} and $N_{AL}(N_{AH} + N_{AL} =$

N_A), where H denotes the group with higher environmental preference. We assume in this section that commuters with higher environmental preferences travel in cleaner cars with a lower value of emission parameters (λ_{fH} and λ_{qH}) than group L has (λ_{fL} and λ_{qL}).

Under nudging information, the perceived travel prices of these two groups are:

$$P_{AH}(t_d) = \mu_{eH} e_{AH}(t_d) + (\alpha + \lambda_h)\left(T_{Af} + q_A(t_d)\right) + \begin{cases} \beta(t^* - T_{Af} - t_d - q_A(t_d)) & t^* \geq T_{Af} + t_d + q_A(t_d) \\ \gamma(T_{Af} + t_d + q_A(t_d) - t^*) & t^* \leq T_{Af} + t_d + q_A(t_d) \end{cases} \tag{45}$$

$$P_{AL}(t_d) = \mu_{eL} e_{AL}(t_d) + (\alpha + \lambda_h)\left(T_{Af} + q_A(t_d)\right) + \begin{cases} \beta(t^* - T_{Af} - t_d - q_A(t_d)) & t^* \geq T_{Af} + t_d + q_A(t_d) \\ \gamma(T_{Af} + t_d + q_A(t_d) - t^*) & t^* \leq T_{Af} + t_d + q_A(t_d) \end{cases} \tag{46}$$

where $P_{AH}(t_d)$ denotes the travel price of road commuters with higher environmental preferences and $P_{AL}(t_d)$ for the lower environmental preference group.

Based on the generalized travel prices in Section 4.1, we find that the relative values of $\mu_{eH}\lambda_{qH}$ and $\mu_{eL}\lambda_{qL}$ determine whether these two groups travel together, and if not in which order they arrive. Since the relative sizes can go either way, the travel pattern is not generally predetermined in this case. When $\mu_{eH}\lambda_{qH} > \mu_{eL}\lambda_{qL}$, the group with the higher environmental preferences and cleaner cars will travel at the shoulder of the peak, while the lower group will travel at the center of the peak. When $\mu_{eH}\lambda_{qH} < \mu_{eL}\lambda_{qL}$, the opposite occurs. Note that in the former case, the more polluting cars will drive at moments where their emissions per trip are higher.

4.3 Heterogenous income and car types

One may also expect the type of car to be related to the driver's income. Hence, in this subsection, we further explore the case of heterogeneous commuters with different incomes and car types. We again consider two discrete groups now, with different incomes. Assuming otherwise identical preferences, but income to affect the marginal utility of income, this means that $\alpha, \beta, \gamma, \lambda_h$ and g vary proportionally between the two groups of travelers, so that the ratios $\frac{\alpha}{\beta}, \frac{\alpha}{\gamma}, \frac{\alpha}{\lambda_h}$ and $\frac{\alpha}{g}$ are constants.[2] The equilibrium numbers of commuters in these two groups are again denoted by N_{AH} and $N_{AL}(N_{AH} + N_{AL} = N_A)$, where H now denotes the group with higher income.

Since cleaner energy vehicles are generally more expensive than diesel vehicles, we could assume that commuters with higher income are more likely to buy cleaner cars with relatively lower emission parameters (λ_{fH} and λ_{qH}) than the lower income groups. This leads to:

Assumption a: *Commuters with higher income travel in relatively cleaner cars than the lower income group.*

When there are no incentives, two groups will travel jointly since they have the same equilibrium travel time development over time ($\frac{\beta_L}{\alpha_L + \mu_h\lambda_{hL}} = \frac{\beta_H}{\alpha_H + \mu_h\lambda_{hH}}$). The generalized travel price in user equilibrium can then be derived as:

$$P_{AL} = (\alpha_L + \mu_h\lambda_{hL} + \mu_e\lambda_{fL})T_{Af} + \frac{\beta_L\gamma_H}{\beta_L+\gamma_L}\frac{N_{AL}+N_{AH}}{s} \tag{47}$$

$$P_{AH} = (\alpha_H + \mu_h\lambda_{hH} + \mu_e\lambda_{fH})T_{Af} + \frac{\beta_H\gamma_H}{\beta_H+\gamma_H}\frac{N_{AL}+N_{AH}}{s} \tag{48}$$

where P_{AL} is the travel price of auto commuters with lower income, and P_{AH} denotes the travel price of auto commuters with higher income.

Since we can derive that $\frac{\beta_L}{\alpha_L + \lambda_{hL} + \mu_e\lambda_{qL}} < \frac{\beta_H}{\alpha_H + \lambda_{hH} + \mu_e\lambda_{qH}}$, there will be sorting under nudging information: the group with the higher income and cleaner cars will travel during the center of the peak, while the group with lower income and more polluting car will travel during the shoulder. The generalized travel price under nudging information can then be derived as:

$$P_{AL}^T = (\alpha_L + \lambda_{hL} + \mu_e\lambda_{fL})T_{Af} + \frac{\beta_L\gamma_L}{\beta_L+\gamma_L}\frac{N_{AL}+N_{AH}}{s} \tag{49}$$

---

[2] We assume that income effects are uniform across travelers' value of time, schedule delays, crowding, and health considerations.

$$P_{AH}^{T} = (\alpha_H + \lambda_{hH} + \mu_e \lambda_{fH}) T_{Af} + \frac{\beta_H \gamma_H}{\beta_H + \gamma_H} \frac{\frac{\beta_L(\alpha_H + \mu_e \lambda_{qH} + \lambda_{hH})}{\beta_H(\alpha_L + \mu_e \lambda_{qL} + \lambda_{hL})} N_{AL} + N_{AH}}{s} \tag{50}$$

However, an alternative to Assumption a may also be plausible, namely when we consider the case where the car's emission parameter depends primarily on the size of car. Higher-income group can then be expected to have bigger cars with higher emission parameters (λ_{fH} and λ_{qH}), than the lower-income. Then, an alternative Assumption *b* can be made:

Assumption b: *Commuters with higher income travel in relatively more polluting than the lower income group.*

The generalized travel cost and departure pattern in the base equilibrium without nudging is then the same under Assumption b as under Assumption a.

However, different from the situation under Assumption a, the relative sizes of $\frac{\beta_L}{\alpha_L + \lambda_{hL} + \mu_e \lambda_{qL}}$ and $\frac{\beta_H}{\alpha_H + \lambda_{hH} + \mu_e \lambda_{qH}}$ are ambiguous under Assumption b. The departure pattern of road commuters under nudging information can characterized as: when $\frac{\beta_L}{\alpha_L + \lambda_{hL} + \mu_e \lambda_{qL}} < \frac{\beta_H}{\alpha_H + \lambda_{hH} + \mu_e \lambda_{qH}}$, the group with the higher income travels during the center of the peak, and the other group in the shoulder of the peak. When $\frac{\beta_L}{\alpha_L + \lambda_{hL} + \mu_e \lambda_{qL}} > \frac{\beta_H}{\alpha_H + \lambda_{hH} + \mu_e \lambda_{qH}}$, the opposite occurs.

4.4 Heterogenous income and environmental preferences

A next potentially relevant relationship may exist between auto commuters' income and environmental preferences. For simplicity, again two discrete groups are considered, and commuters with higher income are now assumed to be more concerned with the environment, thus having a stronger environmental preference.

Without incentives, there is proportional heterogeneity, and the two groups will travel together. The generalized travel prices in user equilibrium then are obtained to:

$$P_{AL} = (\alpha_L + \mu_h \lambda_{hL}) T_{Af} + \frac{\beta_L \gamma_H}{\beta_L + \gamma_L} \frac{N_{AL} + N_{AH}}{s} \tag{51}$$

$$P_{AH} = (\alpha_H + \mu_h \lambda_{hH}) T_{Af} + \frac{\beta_H \gamma_H}{\beta_H + \gamma_H} \frac{N_{AL} + N_{AH}}{s} \tag{52}$$

where P_{AL} gives the travel price for auto commuters with the lower income, and P_{AH} for the others.

Under nudging information incentives, since the relative values of $\frac{\beta_L}{\alpha_L + \lambda_{hL} + \mu_{eL}\lambda_q}$ and $\frac{\beta_H}{\alpha_H + \lambda_{hH} + \mu_{eH}\lambda_q}$ can go either way, we again distinguish between two cases. When $\frac{\beta_L}{\alpha_L + \lambda_{hL} + \mu_{eL}\lambda_q} < \frac{\beta_H}{\alpha_H + \lambda_{hH} + \mu_{eH}\lambda_q}$, the group with higher income and environmental preferences travels during the center of the peak, and the other group travels during the shoulder. When $\frac{\beta_L}{\alpha_L + \lambda_{hL} + \mu_{eL}\lambda_q} > \frac{\beta_H}{\alpha_H + \lambda_{hH} + \mu_{eH}\lambda_q}$, the opposite departure order applies.

Under the above different assumptions on commuter heterogeneity, the effect of nudging information is summarized in Proposition 4. We will use numerical analysis to further explore the effects of nudging information.

Proposition 4. *Nudging information is more likely to produces a negative effect on social surplus especially in two scenarios:*

(a) when crowding effects in the metro system are larger

(b)when welfare-reducing departure time swaps occur between heterogenous road travel groups, due to nudging information.

Proof. Proof of Proposition 4 is provided in Appendix E.

Proposition 4 suggests that, also when crowding effects in the metro system are smaller, the nudging information can still possibly harm the social welfare by causing welfare-reducing swaps in arrival patterns, notably when shifting more polluting cars to the center of peak, where emissions per trip are higher due to longer travel delays.

5. Numerical Analyzes

5.1 Calibration of the numerical models

We use numerical analysis to illustrate the effects of nudging information for heterogeneous commuters. The values of the parameters used in the numerical analysis are in Table 1.

Table 1. Parameters used in the numerical analysis

Parameter	Definition	Value
a_{AL}	Intercept of inverse demand for auto commuters in lower group	55.06
a_{AH}	Intercept of inverse demand for auto commuters in higher group	55.06
a_{ML}	Intercept of inverse demand for metro commuters in lower group	36.98
a_{MH}	Intercept of inverse demand for metro commuters in higher group	36.98
b_{AL}	Slope of inverse demand for auto commuters in lower group	-0.007
b_{AH}	Slope of inverse demand for auto commuters in higher group	-0.007
b_{ML}	Slope of inverse demand for metro commuters in lower group	-0.004
b_{MH}	Slope of inverse demand for metro commuters in higher group	-0.004
η_L	The cross-price effect in lower group	-0.001
η_H	The cross-price effect in higher group	-0.001
α	Value of travel time	10
β	Value of early arrival	6.09
γ	Value of late arrival	23.77
h	Time interval between successive trains	0.05
λ_f	Emission cost per unit time spent on cruising	0.796
λ_q	Emission cost per unit time spent on queuing	1.282
T_{Mf}	Fixed travel time for a metro commuter	1/5
T_{Af}	Uncongested travel time for a road commuter	1/6
g	Value of body congestion	4
λ_h	Extra value of in-vehicle travel time respect to health	2
μ_{eL}	Environmental preference for lower group	0.2
μ_{eH}	Environmental preference for higher group	0.5
μ_h	Perceived health cost parameter without information	0.8
M	Number of total service run	60
M^*	Desired service run with no schedule cost	40
c	Capacity of a train	500
s	Capacity of road bottleneck	3000
σ	Unweighted average scheduling cost for trains	8.12

Following Yu et al. (2023), we set the value of time, the value of schedule delay early, and the value of schedule delay late, as $\alpha = 10, \beta = 6.09$, and $\gamma = 23.77$, respectively. The currency considered is the Euro. And, based on the value of emission parameters used in Liu et al. (2015) and the value of the crowding congestion parameter used in de Palma et al. (2017), we set $\lambda_q$and λ_f at 1.392 and 0.864 per hour, and g at 4. In addition, as health-related costs are usually ignored due to their lesser significance compared to travel time costs, to ensure the ratio of these two types of costs remains within a reasonable range (namely: $0 < \frac{\lambda_h}{\alpha} < 1$), the health parameter λ_h is assumed to be 2. In the base case, we set the supply-side parameters M, M^*, and h as 60, 40, 0.05 *hour,* respectively. The uncongested travel time in the road and metro system is considered as 1/5 hour (12 minutes), 1/6 hour (10 minutes), respectively. The capacity of the road bottleneck is set at 3,000 cars per hour in the road bottleneck, and the capacity of each service run in the metro system is 500, indicating that the same 3-hour travel period are considered in both road and metro systems. The base equilibrium numbers of travelers with high and low environmental preferences in road commuting group is set at 4,500 each, while metro commuting is chosen by 6000 travelers in each group. To accomplish the above goals, a_{AL} =55.06, a_{AH} =55.06, a_{MH} =36.98, $a_{ML} = 36.98$, $b_{AL} = -0.007$, $b_{AH} = -0.007$, $b_{ML} = -0.004$, $b_{MH} = -0.004$, $\eta_L = -0.001$,$\eta_H = -0.001$ are calibrated (Note that these are the rounded values of the above parameter in the inverse demand function). Scenario (a) in Proposition 4 will not occur in the following numerical analysis, since the crowding effect is sufficiently small in our base calibration and sensitivity analyses.

Table 2. Basic equilibrium

Number of car commuters in each group	4500
Number of metro commuters in each group	6000
Mode share-auto	42.86%
Mode share-metro	57.14%
Demand elasticity-auto with respect to generalized price (each group)	-0.50
Demand elasticity-metro with respect to generalized price (each group)	-0.50
Cross demand elasticity-metro to car (each group)	0.10

We begin by examining the effects of nudging information when commuters are heterogeneous and how these impacts change with the degree of heterogeneity between the two groups. To that end, we vary μ_{eH} and μ_{eL} such that the average remains 0.5. Let SW and SW_I denote the equilibrium social welfare without and with nudging information, respectively. SW_T and SW_{IT} denote the social welfare under second-best flat road toll without and with nudging information. Furthermore, SW_{FB} represents the equilibrium under first-best pricing. Note that first-best pricing includes dynamic tolls in both metro and road bottlenecks to optimize departure time patterns; and this ensures that perfectly informed commuters' generalized travel costs equal the marginal travel costs. We define the index $W_I = \frac{SW_I - SW}{SW_{FB} - SW}$ to represent the relative welfare improvement produced by nudging information, compared to that achieved under first-best pricing, using the equilibrium without incentives as the benchmark. Similarly, $W_T = \frac{SW_T - SW}{SW_{FB} - SW}$ and $W_{IT} = \frac{SW_T - SW}{SW_{FB} - SW}$ denote the relative social welfare improvements generated by the second-best flat road toll and the combination of nudging information and the second-best flat road toll, respectively, again compared to the social welfare improvement achieved under first-best pricing.

5.2 Cases without welfare-reducing swaps of departure time due to nudging information

First, we consider cases where nudging information does not lead departure time swaps between the two groups, under different assumptions of heterogenous commuters.

5.2.1 Case I: μ_e varies between groups

We initially discuss the relatively simple case where commuters are heterogenous in environmental preferences, while the commuters with higher environmental preference travel during the center while the lower group travel during the shoulder.

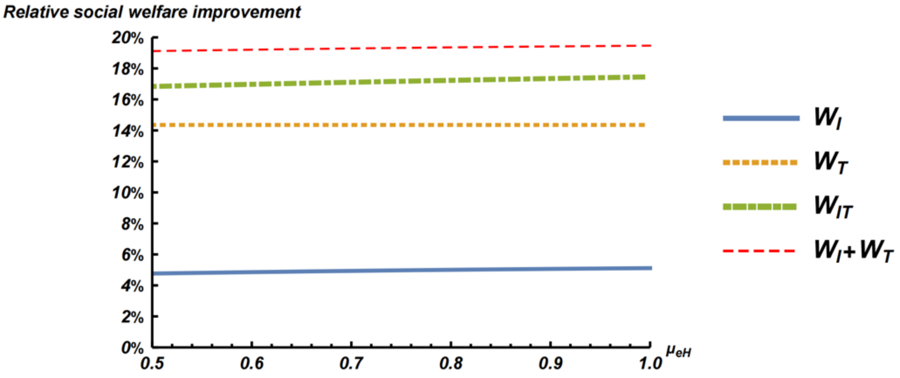


Figure 2. Relative welfare improvement by different policies under varying $\mu_{eH}(\frac{\mu_{eH}+\mu_{eL}}{2} = 0.5)$

Figure 2 mainly shows the relative performance of three different policies compared to first-best tolling and the sum of relative welfare gain from nudging information and flat tolling. Initially, as $W_I > 0$, nudging information can enhance social welfare by encouraging road commuters to internalize health and emission costs during travel, and more so when heterogeneity in environmental preferences between these two groups becomes more pronounced. The positive effects of internalizing emission and health costs within the two groups in this case outweigh the negative impacts of shifting commuters with relatively higher uninternalized emission externalities to the central period. Consequently, total social welfare improvement increases, albeit mildly, when the divergence in environmental preferences between the two groups grows (Similar results can be found in Figure F.1 and F.2). Note that the benefits from nudging information are modest around 5% of benefits from optimal pricing, while the second-best flat road toll can generate a relatively higher welfare improvement of around 14%. Since $W_{IT} > W_T$, nudging information serves as a complement to the second-best flat road toll. As $W_{IT} < W_I + W_T$, the benefits of information and flat tolling are subaddtive. Finally, Figure 2 shows that the heterogeneity in environmental preferences has a small impact on the relative performance of the instruments: the cures are remarkably flat. Results are thus robust with respect to the ratio of μ_{eH} to μ_{eL}.

5.2.2 Case II: λ_f and λ_q vary with μ_e between groups

We now incorporate additional commuter heterogeneity in car types into the previous analysis. When road commuters with higher environmental preferences drive cleaner vehicles (recall the discussion in Section 4.2), λ_f and λ_q decrease with μ_e. We let λ_f and λ_q increase at a decreasing rate when μ_e falls, by setting it in proportion with $(1-\mu_e)^2$. It then holds that $\lambda_{qL}\mu_{eL} > \lambda_{qH}\mu_{eH}$. As a result, with nudging information, commuters with higher environmental preferences travel during the central peak period, while those with lower environmental preferences travel during the shoulder.

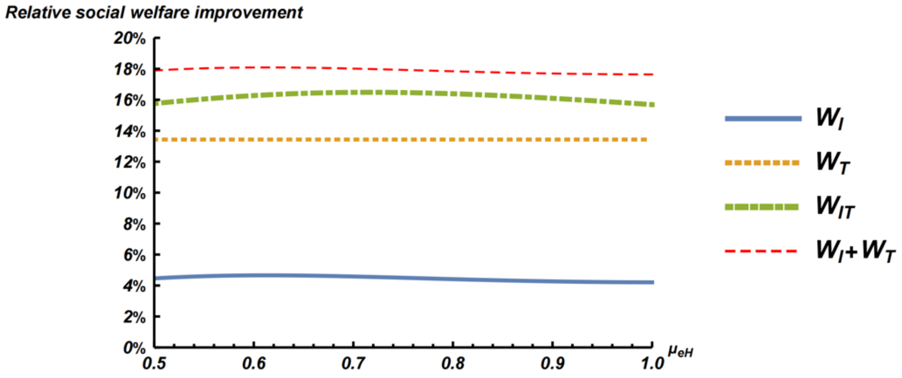


Figure 3. Relative welfare improvement by different policies under varying $\mu_{eH}(\frac{\mu_{eH}+\mu_{eL}}{2} = 0.5)$

The order of four types of relative welfare gains in Figure 2 remains consistent in Figure 3. As shown in Figure 3, since information can shift cleaner cars to the central peak period, the welfare improvement can increase when the differences in perceived emission costs between the two groups become more pronounced. In this scenario, information still serves as a complement to the second-best flat road toll, thus boosting the overall welfare effect ($W_{IT} > W_T$). When raising μ_{eH}, the social welfare gained from information, and from the combination of information and the second-best flat road toll, both first increases and then decreases, while it remains positive. This occurs because the difference in the ratio of perceived emission costs $\mu_e\lambda_q$ between two groups due to nudging becomes more pronounced at first and then diminishes when μ_{eH} increases, particularly when λ_f and λ_q both vary in proportion $(1-\mu_e)^2$ as assumed. Since $W_{IT} < W_I + W_T$, the combined benefits of nudging information and flat tolling remain subadditive in this scenario. Again, the relative efficiency scores remain rathe constant in Figure 3.

5.2.3 Case III: λ_f and λ_q vary with α between groups

Different patterns of relatively welfare improvements emerge when considering commuter incomes. We now extend the analysis to incorporate heterogeneity in both commuter income and car types. To consider where the case $\frac{\beta_L}{\alpha_L + \lambda_{hL} + \mu_e\lambda_{qL}} > \frac{\beta_H}{\alpha_H + \lambda_{hH} + \mu_e\lambda_{qH}}$, so that the higher-income group travels during the shoulder of the peak with information, we assume that $\lambda_f^{\frac{1}{2}}$ and $\lambda_q^{\frac{1}{2}}$ both vary in proportion with α. This implies the λ_f and λ_q vary stronger than proportionally with α.

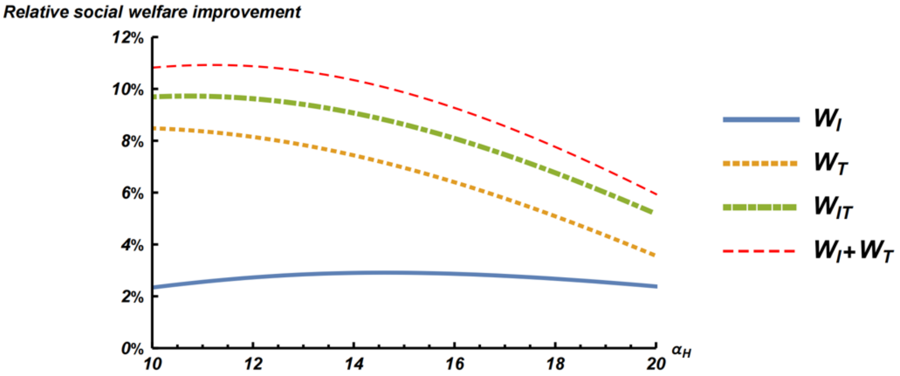


Figure 4. Relative welfare improvement by different policies under varying $\alpha_H(\frac{\alpha_L+\alpha_H}{2}=10)$

The order of relative welfare improvement remains the same in Figure 4. Higher-income groups with more polluting cars now travel during the shoulder of the peak, while lower-income groups with cleaner cars travel during the central peak period with nudging information. Information incentives again enhance welfare levels and serve as a complement to the second-best flat road toll ($W_{IT} > W_T$); however, we also see that the impacts of information alone are more modest, and peak around $\alpha_H = 15$. The reduction in gains for further increase of α_H occurs because of the negative effect of moving the higher-scheduling group, with higher schedule delay cost, to the shoulder of the peak, this then outweighs the positive effect of shifting the cleaner cars to the central peak and further lead to the U shape of W_I, W_{IT}. W_T decreases monotonically with α_H. This pattern arises because the efficiency of undifferentiated tolling diminishes when heterogeneity between commuter groups become more significant. The combined benefits of nudging information and flat tolling remain subadditive in this scenario. A similar case is provided in Appendix F (Figure F.3).

5.3 Cases with welfare-reducing swaps of departure time by nudging information

In this subsection, we focus on the cases where nudging information may lead to welfare-reducing swaps of departure times between two groups.

5.3.1 Case IV- λ_f and λ_q vary with μ_e between groups

Recall the relationship between commuters' environmental preferences and car types we have discussed in section 4.2. For example, when λ_f and λ_q both vary in proportion $(1-\mu_e)^{\frac{1}{2}}$ reflecting that a higher environmental preference is associated with cleaner cars, we can obtain a case where $\lambda_{qL}\mu_{eL} < \lambda_{qH}\mu_{eH}$. With nudging information, in the equilibrium, commuters with lower environmental preferences and more polluting cars may then travel during the central peak period, while those with higher environmental preferences and cleaner cars travel during the shoulder, while without nudging information, the two commuting groups were traveling together.

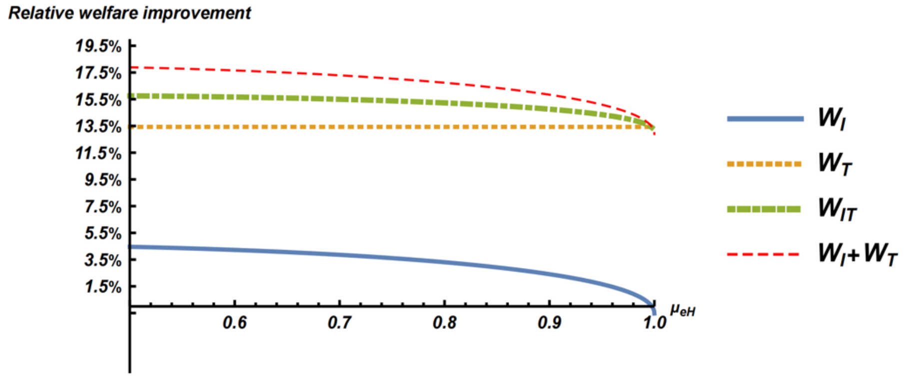


Figure 5. Relative welfare improvement by different policies under varying μ_{eH} ($\frac{\mu_{eH}+\mu_{eL}}{2}=0.5$)

Information then shifts more polluting cars to the central peak period, while cleaner cars are pushed to the shoulder of the peak. This could still be a welfare-improving change when the unit cost of emission externality is not too different between the two groups, while scheduling and travel delay costs support this shift. However, as shown in Figure 5, when the difference in unit emission cost between the two groups becomes large enough (with μ_{eH} approaching 1), nudging information may eventually have a negative effect on social welfare as it increases total emissions. It is important to note that when nudging information generates a negative effect, it also no longer complements the second-best flat road toll (W_{IT} becomes smaller than W_T). Furthermore, in such cases, the combined benefits of nudging information and flat tolling become superadditive ($W_{IT} > W_I + W_T$). This outcome aligns with the intuition behind second-best paradoxes, where measures that are beneficial under optimal conditions—such as providing perfect information—may produce adverse effects in second-best scenarios, such as when no road pricing is implemented. Again, similar cases may occur for information provision on uncertain travel time (Verhoef et al., 1996a).

5.3.2 Case V- λ_f and λ_q vary with α between groups

Next, we consider the case when ${\lambda_f}^2$ and ${\lambda_q}^2$ both vary in proportion α. Recall that Assumption b in Section 4.3 stipulates that high-income travelers drive more polluting cars. From this, we can derive that $\frac{\beta_L}{\alpha_L + \lambda_{hL}+\mu_e\lambda_{qL}} < \frac{\beta_H}{\alpha_H+\lambda_{hH}+\mu_e\lambda_{qH}}$, which indicates that with information provision the higher-income group would travel during the central peak period, while the lower-income group would travel during the shoulder of the peak.

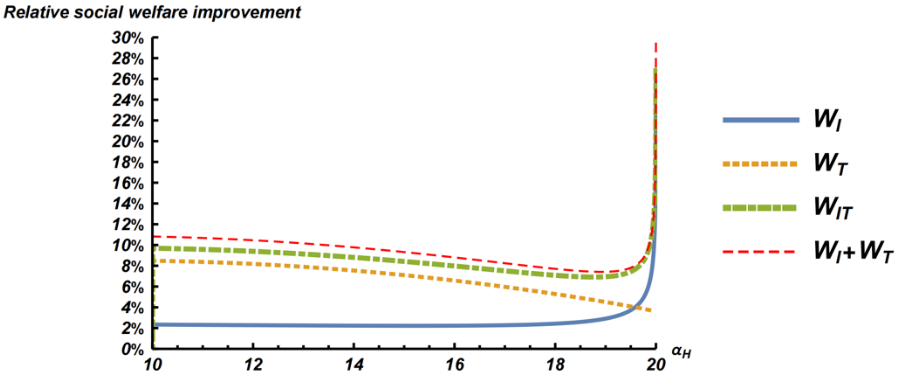


Figure 6. Relative welfare improvement by different policies under varying $\alpha_H(\frac{\alpha_L+\alpha_H}{2} = 10)$

In Figure 6, higher-income groups with more polluting cars travel during the central peak period, while lower-income groups with cleaner cars travel during the shoulder of the peak when nudging information is applied. This may still increase welfare and also complement to the second-best flat road toll when it encourages individuals to internalize emission externalities and unperceived health costs during driving, on top of shifting commuters with higher schedule costs to the central peak period. However, welfare gains are relatively smaller than in the other cases because more polluting cars are being shifted to the central peak. No super additivity between nudging information and flat tolling arises in this case. Note that information leads to progressively high relative efficiency when α increases. This occurs because the temporal separation brings larger benefits to the high-income drivers. Also, their schedule delay cost parameters are high. When information pushes low income to the shoulders, high income drivers have higher schedule delay cost advantages.

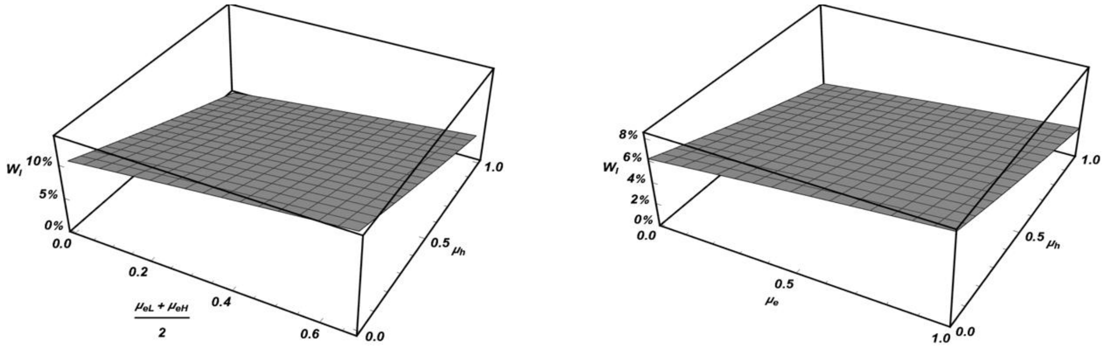
A further key difference with the previous scenario is that the welfare improvement from nudging information surpasses that of the second-best flat toll when α_H approaches 20. This indicates that the flat road toll scheme may become less efficient compared to nudging information when the heterogeneity in commuters' income becomes significant (similar results are shown in Figures F.1 and F.2).

6. Robustness check of relative welfare improvement

In the previous section, implementing nudging information has been shown to effectively enhance social welfare in most instances under the majority of parameter values considered. Furthermore, in these scenarios, the order of relative welfare improvements considered (W_I, W_T, W_{IT} and $W_I + W_T$) remains stable when no welfare-reducing swap occur. However, only when a welfare-reducing swap in arrival times occurs, nudging information lowers welfare.

To further test the sensitivity of our results, we change some other key parameters. Specifically, we review the cases discussed in the last section to test the robustness with respect to the emission and health parameters.

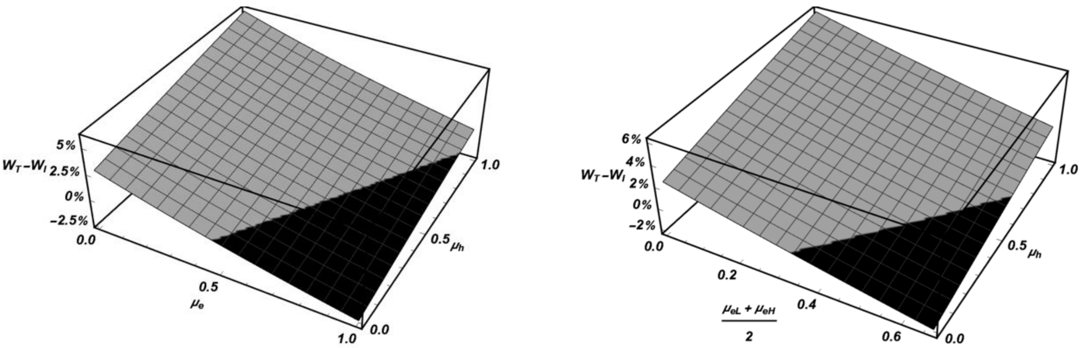
We first investigate the robustness of positive effect produced by nudging information ($W_I > 0$) when no welfare-reducing swap occurs. We select two typical positive scenarios where the relative welfare improvement is either the highest (around 10%) or the lowest (around 6%) among all positive cases, as illustrated in Figures 7(a) and 7(b), respectively. (The remaining positive cases are listed in Appendix G.)



(a) μ_e varies between groups (b) $\lambda_f^{\frac{1}{2}}$ and $\lambda_q^{\frac{1}{2}}$ both vary in proportion α

Figure 7. Relative welfare improvement by nudging information under varying μ_e and $\mu_h(\mu_{eH} = 3\mu_{eL})$

Since the order $W_I < W_T < W_{IT} < W_I + W_T$ is relatively stable for cases without welfare-reducing swaps, we focus exclusively on cases where $W_I > W_T$ arises. As illustrated in Figures 8(a) and 8(b), by considering the heterogeneity in income, the effect of nudging information can surpass the impact of second-best flat tolls, particularly when commuters exhibit a higher degree of environmental consciousness (higher μ_e). In such scenarios, nudging can exert a relatively greater impact by causing positive departure time swaps, thereby enhancing relatively higher welfare outcomes than flat tolling. Note that in the following figures, the black segments represent negative values.



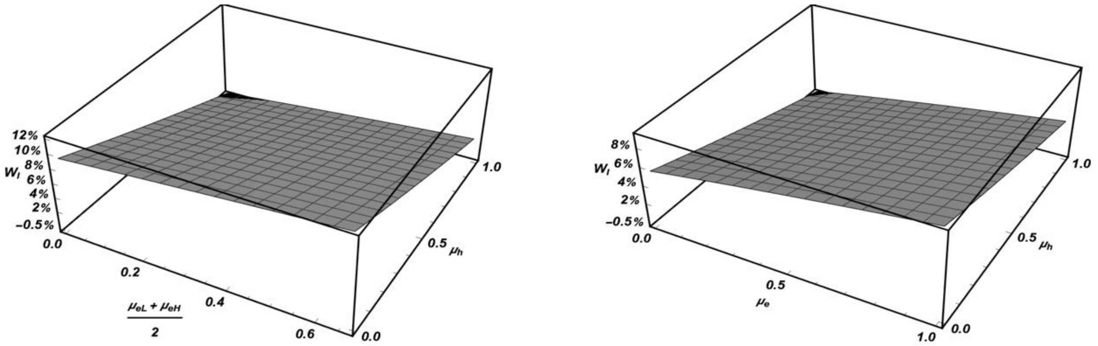
(a) $\frac{1}{\lambda_q}$ and $\frac{1}{\lambda_f}$ vary in proportion α (b) ${\mu_e}^2$ varies in proportion α

Figure 8. Relative welfare improvement by nudging under varying μ_e and $\mu_h(\mu_{eH} = 3\mu_{eL})$

We next explore the specific circumstances under which the negative effects of nudging information arise, and examine the impact on the relative order of

W_I , W_T, W_{IT} and $W_I + W_T$.

Figures 9(a) and 9(b) show that a negative welfare effect of nudging information is more likely to happen when the two groups exhibit near indifference toward the environment and internalize most of the health costs associated with travel in the absence of information. In such scenarios, the more polluting group can impose additional emission costs, ultimately dominating the overall welfare impacts and compromising the entire outcome.



(a) λ_f and λ_q vary in proportion $(1-\mu_e)^{\frac{1}{2}}$

(b) $\lambda_f{}^2$ and $\lambda_q{}^2$ vary in proportion α

Figure 9. Relative welfare improvement by nudging under varying μ_e and $\mu_h(\mu_{eH} = 3\mu_{eL})$

In such cases, where nudging information leads to negative welfare effects, we can also find a different order of W_I , W_T, W_{IT} and $W_I + W_T$, as shown in more detail Appendix G(Figure G.2(b), Figure G.3(a) and 3(b), Figure G.4(a) and 4(b)). These negative effects arise primarily due to welfare-reducing swaps of arrival times, which prevent nudging information from consistently serving as a complement to the second-best flat road toll($W_T < W_{IT}$). In these scenarios, the combined benefits of nudging information and flat tolling typically exhibit super additivity: $W_{IT} > W_I + W_T$. Examining the relationship between W_I and W_T, it becomes obvious that welfare-reducing swaps are not the primary cause of $W_I > W_T$, as illustrated in Figures Appendix G (Figure G.2(a) and 2(b)). Instead, the direct factors contributing to the relatively higher welfare improvement from nudging information are the heterogeneity in commuters' income and their environmental preferences, which is consistent to the discussion around Figure 8(a) and 8(b).

7. Conclusion

We investigated the impacts of providing “nudging” information on managing congestion and emissions in a bi-modal road and metro network. By incorporating information about health and emission costs, commuters can be encouraged to internalize these externalities, leading to different travel patterns, and potentially improving social welfare.

The research highlights the potential effectiveness of nudging information as a policy tool. By making individuals aware of the health risks and environmental impact of their travel choices, it is possible to encourage more sustainable and health-conscious decisions. This awareness can reduce the negative externalities associated with congestion and emissions. For instance, when commuters are informed about the adverse health effects for others of exposure to vehicle emissions, and the own increased risk of respiratory and cardiovascular diseases, they may opt for cleaner and more sustainable travel modes. In our numerical examples, we typically find modest positive welfare effects that amount to less than 10% of the welfare gains from optimal pricing. Higher gains occur when all impacts are positive: information guides lower income groups with the more polluting cars to times where they emit less (i.e., the shoulders of the peak), and impacts on scheduling and travel delay costs are positive. Negative effects may arise in opposite cases, in particular when the costs of emissions are high but (largely) ignored by drivers who are pushed towards the central peak when they are informed.

Under the assumption that commuters are homogeneous, meaning they have the same preferences and respond identically to information, we find that providing nudging information uniformly generally improves social welfare except when the crowding externalities are excessively large in the metro system. When all commuters consider the health and emission costs of their trips, they are more likely to avoid congested routes or peak travel times, reducing overall congestion and associated emissions. This shift not only alleviates traffic bottlenecks but also promotes a more even distribution of traffic across periods and modes, enhancing the efficiency of the transportation network. We further prove that in the cases where a naive flat Pigouvian road toll improves total social welfare, nudging information always serves as a welfare-improving strategy by effectively reducing the negative externalities of congestion and emissions. Moreover, regardless of the extreme crowding externalities in the metro system, nudging information always serves as a complement to the second-best flat road toll in allowing an increase in welfare gains.

The study also explores the more complex scenario where commuters are heterogeneous, meaning they have diverse preferences, behaviors, and sensitivities to nudging information. In this case, the effectiveness of nudging information may vary significantly and in complex ways with patterns of heterogeneity. Commuters with different environmental preferences, car types, and income levels respond differently to the provided information. For example, high-income commuters with a strong environmental preference may readily shift to less polluting travel options when informed about emission costs, while low-income commuters with less concern for environmental issues may be less responsive to the same information. But higher incomes may have cars that are more, or less, polluting than lower incomes, further complicating the overall effects of nudging information on emissions and social welfare.

When commuters are heterogeneous, the implementation of nudging leads to varied travel patterns, potentially reducing congestion and emissions in most cases. Also without excessively large crowding externalities in the metro system, there may be other scenarios where nudging may inadvertently lead to negative welfare outcomes. For instance, if commuters are largely indifferent to environmental issues, the information may have little to no effect. Furthermore, if the nudging leads to more polluting vehicles traveling during peak periods, it could exacerbate congestion and emissions, resulting in even more severe externality problems. In these cases, nudging information cannot always serve as a complement to the second-best flat road toll. This highlights the importance of understanding the specific characteristics of the commuter population when designing and implementing nudging strategies.

The findings of this study have significant policy implications for urban planners and policymakers. They suggest that a nuanced approach to information dissemination, which accounts for the diverse characteristics of commuters, can be more effective in achieving the desired outcomes. For instance, targeted campaigns that address specific concerns of different commuter groups, such as health risks for those more sensitive to pollution or cost savings for those with tighter budgets, can enhance the effectiveness of nudging information.

While the study provides valuable insights into the potential benefits of nudging information, we also acknowledge several challenges and limitations. One significant challenge is accurately predicting and measuring the behavioral responses of commuters to nudging information, especially in a heterogeneous population. Differences in cultural, social, and economic contexts can influence how commuters perceive and react to information, making it difficult to generalize findings across different settings. It would be interesting to empirically test the predictions and assumptions of the paper and to test if the effect differs between the short run and long run. This uncertainty in parameter values is also one of the reasons we did extensive sensitivity checks.

Another limitation is the potential for unintended consequences. As the study suggests, nudging information may not always lead to positive outcomes, especially if it encourages behavior that increases congestion or emissions. Policymakers must carefully design and monitor nudging strategies to mitigate such risks, and ensure that the overall impact is positive.

Future research could explore the role of technology in enhancing the effectiveness of nudging information. With the increasing availability of real-time data and advanced analytics, there are opportunities to develop more personalized and dynamic nudging strategies that adapt to the changing needs and preferences of commuters. Exploring these possibilities can help policymakers leverage technology to create more responsive and efficient transportation systems.

In conclusion, this study demonstrates the potential of nudging information as a valuable tool for managing congestion and emissions in urban transportation networks while showing that the overall impacts will most likely fall short of those of price policies. However, it may be much easier politically to implement nudging than pricing, and it is also a much cheaper option. The effectiveness of nudging strategies depends on a deep understanding of the heterogeneity among commuters and careful consideration of the potential for unintended consequences. As urban populations continue to grow, and environmental concerns become ever more pressing, innovative approaches like nudging can play an ever-important role in urban mobility.

Acknowledgment

This study is supported by the National Natural Science Foundation of China (72071017) and (72474018), and Chinese Scholarship Council (CSC) Scholarship (202307090005). Any remaining errors are solely our responsibility.

References

[1] Arnott, R., de Palma, A., & Lindsey, R. (1990). Departure time and route choice for the morning commute. Transportation Research Part B: Methodological, 24(3), 209-228.

[2] Bhargava, S., & Loewenstein, G. (2015). Behavioral economics and public policy 102: Beyond nudging. American Economic Review, 105(5), 396-401.

[3] Carroll, P., Caulfield, B., & Ahern, A. (2019). Measuring the potential emission reductions from a shift towards public transport. Transportation Research Part D: Transport and Environment, 73, 338-351.

[4] Coria, J., & Zhang, X. B. (2017). Optimal environmental road pricing and daily commuting patterns. Transportation Research Part B: Methodological, 105, 297-314.

[5] de Palma, A., & Lindsey, R. (1998). Information and usage of congestible facilities under different pricing regimes. Canadian Journal of Economics, 666-692.

[6] de Palma, A., & Lindsey, R. (2011). Traffic congestion pricing methodologies and technologies. Transportation Research Part C: Emerging Technologies, 19(6),

1377-1399.

[7] de Palma, A., Lindsey, R., & Monchambert, G. (2017). The economics of crowding in rail transit. Journal of Urban Economics, 101, 106-122.

[8] Emmerink, R. H., Verhoef, E. T., Nijkamp, P., & Rietveld, P. (1996). Information provision in road transport with elastic demand: A welfare economic approach. Journal of Transport Economics and policy, 117-136.

[9] Erlei, M. (2008). Heterogeneous social preferences. Journal of Economic Behavior & Organization, 65(3-4), 436-457.

[10] Frank, L. D., Andresen, M. A., & Schmid, T. L. (2004). Obesity relationships with community design, physical activity, and time spent in cars. American journal of preventive medicine, 27(2), 87-96.

[11] Guerassimoff, G., & Thomas, J. (2015). Enhancing energy efficiency and technical and marketing tools to change people's habits in the long-term. Energy and Buildings, *104*, 14-24.

[12] Kamas, L., & Preston, A. (2012). Distributive and reciprocal fairness: What can we learn from the heterogeneity of social preferences? Journal of Economic Psychology, 33(3), 538-553.

[13] Li, Z. C., Lam, W. H., & Wong, S. C. (2012). Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city. Transportation Research Part B: Methodological, 46(1), 30-49.

[14] Li, Z. C., & Zhang, L. (2020). The two-mode problem with bottleneck queuing and transit crowding: How should congestion be priced using tolls and fares? Transportation Research Part B: Methodological, 138, 46-76.

[15] Liu, Q., Jiang, R., Liu, R., Zhao, H., & Gao, Z. (2020). Travel cost budget based user equilibrium in a bottleneck model with stochastic capacity. Transportation Research Part B: Methodological, 139, 1-37.

[16] Liu, W., Yin, Y., & Yang, H. (2015). Effectiveness of variable speed limits considering commuters' long-term response. Transportation Research Part B: Methodological, 81, 498-519.

[17] Lucas, P., de Oliveira, A., & Banuri, S. (2014). The effects of group composition and social preference heterogeneity in a public goods game: An agent-based simulation. Journal of Artificial Societies and Social Simulation, 17(3), 5.

[18] Ma, R., Ban, X. J., & Szeto, W. Y. (2017). Emission modeling and pricing on single-destination dynamic traffic networks. Transportation Research Part B: Methodological, 100, 255-283.

[19] Marchiori, D. R., Adriaanse, M. A., & De Ridder, D. T. (2017). Unresolved questions in nudging research: Putting the psychology back in nudging. Social and Personality Psychology Compass, 11(1), e12297.

[20] Mirabel, F., & Reymond, M. (2011). Bottleneck congestion pricing and modal split: Redistribution of toll revenue. Transportation Research Part A: Policy and Practice, 45(1), 18-30.

[21] Myers, E., & Souza, M. (2020). Social comparison nudges without monetary

incentives: Evidence from home energy reports. Journal of Environmental Economics and Management, 101, 102315.

[22] Small, K.A., Verhoef, E.T., & Lindsey, R. (2024). The Economics of Urban Transportation. Routledge, Chapman & Hall.

[23] Sudarshan, A. (2017). Nudges in the marketplace: The response of household electricity consumption to information and monetary incentives. Journal of Economic Behavior & Organization, 134, 320-335.

[24] Sugiyama, T., Wijndaele, K., Koohsari, M. J., Tanamas, S. K., Dunstan, D. W., & Owen, N. (2016). Adverse associations of car time with markers of cardio-metabolic risk. Preventive medicine, 83, 26-30.

[25] Thomas, H., & Serrenho, A. C. (2024). Using different transport modes: An opportunity to reduce UK passenger transport emissions? Transportation Research Part D: Transport and Environment, 126, 103989.

[26] van den Berg, V.A.C., & Verhoef, E. T. (2014). Congestion pricing in a road and rail network with heterogeneous values of time and schedule delay. Transportmetrica A: Transport Science, 10(5), 377-400.

[27] van den Berg, V.A.C., & Verhoef, E. T. (2011). Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay. Journal of Public Economics, 95(7-8), 983-992.

[28] Verhoef, E.T., Emmerink, R.H., Nijkamp, P., & Rietveld, P. (1996a). Information provision, flat and fine congestion tolling and the efficiency of road usage. Regional Science and Urban Economics, 26(5), 505-529.

[29] Verhoef, E.T., Nijkamp, P., & Rietveld, P. (1996b). Second-best congestion pricing: the case of an untolled alternative. Journal of Urban Economics, 40(3), 279-302.

[30] Vickrey, W. S. (1969). Congestion theory and transport investment. The American Economic Review, 59(2), 251-260.

[31] Vosough, S., de Palma, A., & Lindsey, R. (2022). Pricing vehicle emissions and congestion externalities using a dynamic traffic network simulator. Transportation Research Part A: Policy and Practice, 161, 1-24.

[32] Wang, X., Rodríguez, D. A., Sarmiento, O. L., & Guaje, O. (2019). Commute patterns and depression: Evidence from eleven Latin American cities. Journal of Transport & Health, 14, 100607.

[33] Xiao, L. L., Liu, T. L., & Huang, H. J. (2016). On the morning commute problem with carpooling behavior under parking space constraint. Transportation Research Part B: Methodological, 91, 383-407.

[34] Yang, H., & Tang, Y. (2018). Managing rail transit peak-hour congestion with a fare-reward scheme. Transportation Research Part B: Methodological, 110, 122-136.

[35] Yu, X., van den Berg, V. A.C., & Li, Z. C. (2023). Congestion pricing and information provision under uncertainty: Responsive versus habitual pricing. Transportation Research Part E: Logistics and Transportation Review, 175, 103119.

[36] Yu, Y., Han, X., Jia, B., Jiang, R., Gao, Z. Y., & Zhang, H. M. (2021). Is providing inaccurate pre-trip information better than providing no information in the morning commute under stochastic bottleneck capacity? Transportation Research Part C: Emerging Technologies, 126, 103085.

Appendix

Appendix A. Number of commuters in each mode under different incentives

A.1 No incentives

$$N_A^N = \frac{s(cM\eta(\sigma-a_M)+a_A(-g+cMb_M)+cM\alpha\eta T_{Mf}+(g-cMb_M)T_{Af}(\alpha+\lambda_h\mu_h))}{-g\delta+cMs\eta^2+gsb_A+cM(\delta-sb_A)b_M} \tag{A.1}$$

$$N_M^N = \frac{cM(-s\eta a_A+\delta(\sigma-a_M)+s\alpha\eta T_{Af}+\alpha\delta T_{Mf}-sb_A(\sigma-a_M+\alpha T_{Mf})+s\eta T_{Af}\lambda_h\mu_h)}{-g\delta+cMs\eta^2+gsb_A+cM(\delta-sb_A)b_M} \tag{A.2}$$

A.2 Nudging information

$$N_A^I = \frac{s(cM\eta(\sigma-a_M)+a_A(-g+cMb_M)+cM\alpha\eta T_{Mf}+(g-cMb_M)T_{Af}(\alpha+\lambda_h+\lambda_f\mu_e))}{-g\delta+cMs\eta^2+gsb_A+cM(\delta-sb_A)b_M} \tag{A.3}$$

$$N_M^I = \frac{cM(-s\eta a_A+\delta(\sigma-a_M)+s\alpha\eta T_{Af}+\alpha\delta T_{Mf}-sb_A(\sigma-a_M+\alpha T_{Mf})+s\eta T_{Af}\lambda_h+s\eta T_{Af}\lambda_f\mu_e)}{-g\delta+cMs\eta^2+gsb_A+cM(\delta-sb_A)b_M} \tag{A.4}$$

A.3 Optimal uniform toll

$$N_A{}^O = (s(-a_A(g-cMb_M)^2+g^2T_{Af}(\alpha+\lambda_f+\lambda_h)+c^2M^2b_M^2T_{Af}(\alpha+\lambda_f+\lambda_h)-cMb_M(-cM\eta a_M+cM\eta(\sigma+\alpha T_{Mf})+2gT_{Af}(\alpha+\lambda_f+\lambda_h)))(\alpha+\lambda_h\mu_h))/(sb_A(g-cMb_M)^2(\alpha+\lambda_h\mu_h)+cMb_M(\alpha(4g\delta-cMs\eta^2)+2g\delta(\lambda_h+\lambda_q)+(2g\delta-cMs\eta^2)\lambda_h\mu_h)-g^2\delta(2\alpha+\lambda_q+\lambda_h(1+\mu_h))-c^2M^2\delta b_M^2(2\alpha+\lambda_q+\lambda_h(1+\mu_h))) \tag{A.5}$$

$$N_M{}^O = (cM(-g+cMb_M)(2\alpha\delta\sigma-s\alpha\sigma b_A+s\alpha^2\eta T_{Af}+2\alpha^2\delta T_{Mf}-s\alpha^2b_AT_{Mf}+s\alpha\eta T_{Af}\lambda_f+\delta\sigma\lambda_h+s\alpha\eta T_{Af}\lambda_h+\alpha\delta T_{Mf}\lambda_h+\delta\sigma\lambda_q+\alpha\delta T_{Mf}\lambda_q+\lambda_h(\delta(\sigma+\alpha T_{Mf})-sb_A(\sigma+\alpha T_{Mf})+s\eta T_{Af}(\alpha+\lambda_f+\lambda_h))\mu_h-s\eta a_A(\alpha+\lambda_h\mu_h)+sa_Mb_A(\alpha+\lambda_h\mu_h)-\delta a_M(2\alpha+\lambda_q+\lambda_h(1+\mu_h)))))/(-sb_A(g-cMb_M)^2(\alpha+\lambda_h\mu_h)+cMb_M(-4g\alpha\delta+cMs\alpha\eta^2-2g\delta(\lambda_h+\lambda_q)+(-2g\delta+cMs\eta^2)\lambda_h\mu_h)+g^2\delta(2\alpha+\lambda_q+\lambda_h(1+\mu_h))+c^2M^2\delta b_M^2(2\alpha+\lambda_q+\lambda_h(1+\mu_h))) \tag{A.6}$$

A.4 Optimal uniform toll with nudging information

$$N_A{}^* = -((s(-a_A(g-cMb_M)^2+g^2T_{Af}(\alpha+\lambda_f+\lambda_h)+c^2M^2b_M^2T_{Af}(\alpha+\lambda_f+$$

$$\lambda_h) - cMb_M(-cM\eta a_M + cM\eta(\sigma + \alpha T_{Mf}) + 2gT_{Af}(\alpha + \lambda_f + \lambda_h)))(\alpha + \lambda_h + \lambda_q\mu_e))/\ (-sb_A(g - cMb_M)^2(\alpha + \lambda_h + \lambda_q\mu_e) + g^2\delta(2\alpha + 2\lambda_h + \lambda_q(1 + \mu_e)) + c^2M^2\delta b_M^2(2\alpha + 2\lambda_h + \lambda_q(1 + \mu_e)) + cMb_M(-4g\alpha\delta + cMs\alpha\eta^2 + (-4g\delta + cMs\eta^2)\lambda_h + \lambda_q(-2g\delta + (-2g\delta + cMs\eta^2)\mu_e)))) \tag{A.7}$$

$$N_M{}^* = -((cM(g - cMb_M)(2\alpha\delta\sigma - s\alpha\sigma b_A + s\alpha^2\eta T_{Af} + 2\alpha^2\delta T_{Mf} - s\alpha^2 b_A T_{Mf} + s\alpha\eta T_{Af}\lambda_f + 2\delta\sigma\lambda_h - s\sigma b_A\lambda_h + 2s\alpha\eta T_{Af}\lambda_h + 2\alpha\delta T_{Mf}\lambda_h - s\alpha b_A T_{Mf}\lambda_h + s\eta T_{Af}\lambda_f\lambda_h + s\eta T_{Af}\lambda_h^2 + \delta\sigma\lambda_q + \alpha\delta T_{Mf}\lambda_q + (\delta(\sigma + \alpha T_{Mf}) - sb_A(\sigma + \alpha T_{Mf}) + s\eta T_{Af}(\alpha + \lambda_f + \lambda_h))\lambda_q\mu_e - s\eta a_A(\alpha + \lambda_h + \lambda_q\mu_e) + a_M(sb_A(\alpha + \lambda_h + \lambda_q\mu_e) - \delta(2\alpha + 2\lambda_h + \lambda_q(1 + \mu_e))))))/(-sb_A(g - cMb_M)^2(\alpha + \lambda_h + \lambda_q\mu_e) + g^2\delta(2\alpha + 2\lambda_h + \lambda_q(1 + \mu_e)) + c^2M^2\delta b_M^2(2\alpha + 2\lambda_h + \lambda_q(1 + \mu_e)) + cMb_M(-4g\alpha\delta + cMs\alpha\eta^2 + (-4g\delta + cMs\eta^2)\lambda_h + \lambda_q(-2g\delta + (-2g\delta + cMs\eta^2)\mu_e)))) \tag{A.8}$$

Appendix B. Proof of Proposition 1

The effects of nudging information can be divided into two distinct components based on the generalized road travel price function P_A^I under nudging information:

First, a flat road toll τ_A^I , which is equal to the difference between the fixed travel price with and without nudging information, represented as:

$$\tau_A^I = (\alpha + \lambda_h + \mu_e\lambda_f)T_{Af} - (\alpha + \mu_h\ \lambda_h)T_{Af} = ((1 - \mu_h)\lambda_h + \mu_e\lambda_f)T_{Af} \tag{B.1}$$

Second, a zero value of dynamic road toll, which changes drivers' departure time to internalize the congestion, health and emission cost during queuing at the bottleneck, the internalized external cost TEC_A^I by this type of toll can be written as:

$$TEC_A^I = \mu_e(\lambda_f T_{Af} + \frac{\lambda_q\delta N_{AI}}{2(\alpha+\mu_e\lambda_q+\lambda_h)s})N_{AI} \tag{B.2}$$

Since the metro serves as an imperfect substitution for driving, the flat road toll τ_A^I cannot always generate positive welfare effects due to the crowding in the metro. Specifically, when the marginal cost of shifting drivers to the metro mode is too large (e.g., the value of crowding parameter g is sufficiently big), τ_A^I can generate negative effects.

While the zero-value dynamic toll consistently enhances welfare in the bi-modal system, this positive effect mainly depends on the value of μ_e which is related to environmental preferences. Hence, the possible negative effect of the flat toll τ_A^I is not related to μ_e and is unbounded, and thus it will harm welfare if crowding effects are high enough.

On the contrary, when crowding effects in the metro are relatively minor, the fixed toll τ_A^I is more likely to generate positive welfare effects. In such cases, the zero-value dynamic toll remains welfare-improving, and thus the combined effect of nudging information on total welfare is positive.

This completes the proof.

Appendix C. Proof of Proposition 2

The optimal flat external toll for drivers can be formulated as:

$$\tau_A^{OE} = N_A \frac{\partial P_A}{\partial N_A} + \frac{\partial TE_A}{\partial N_A} + (1-\mu_h)\frac{\partial TH_A}{\partial N_A} = \delta \frac{N_A}{s} + (\lambda_f + (1-\mu_h)\lambda_h)T_{Af} + (\lambda_q + (1-\mu_h)\lambda_h)\frac{\delta N_A}{(\alpha+\mu_h\lambda_h)s} \tag{C.1}$$

The second derivative of total social welfare respect to τ_A can be derived as:

$$\frac{\partial^2 SW}{\partial {\tau_A}^2} = \frac{-(\alpha+\mu_h\lambda_h)\left(c^2M^2s^2\eta^2b_M + s\left((\delta - sb_A)(g-cMb_M)^2\right)\right) - s\delta(g-cMb_M)^2\left((\alpha+\lambda_h)+\lambda_q\right)}{(-g\delta + cMs\eta^2 + gsb_A + cM(\delta - sb_A)b_M)^2(\alpha+\lambda_h\mu_h)} < 0 \tag{C.2}$$

When $\tau_A = -\infty$, the number of road commuters is almost equal to $+\infty$ which indicates that it is beneficial to move commuters from auto to metro. Hence, at this extreme point, we can derive that $\frac{\partial SW}{\partial \tau_A} > 0$.

When $\tau_A = +\infty$, the number of metro commuters is almost equal to $+\infty$ which shows that it is beneficial to move commuters from metro to auto. Hence, at this extreme point, we can derive that $\frac{\partial SW}{\partial \tau_A} < 0$.

Then, we can derive that the welfare first increases and then decreases with flat toll τ_A.

$$\tau_A^I = ((1-\mu_h)\lambda_h + \mu_e\lambda_f)T_{Af} < \tau_A^{OE} \tag{C.3}$$

As discussed in Appendix B, the effect of nudging information can be divided into a fixed toll τ_A^I and zero-value dynamic toll, we can prove that it can always generate welfare improvements in this scenario: as τ_A^{OE} is assumed to always generate positive welfare effects, τ_A^I with a lower value than τ_A^{OE} can also always be a welfare-improving strategy. Moreover, since the change of departure time by a zero value of dynamic toll is always beneficial, we can summarize the total welfare effect of the nudging information is positive.

This completes the proof.

Appendix D. Proof of Proposition 3

The generalized travel price of drivers under a fixed road toll is written as:

$$P_A^T = P_A + \tau_A = (\alpha + \mu_h \lambda_h) T_{\mathrm{Af}} + \delta \left(\frac{N_A}{s} \right) + \tau_A \tag{D.1}$$

The generalized travel price of drivers under both a fixed road toll is written as:

$$P_A^{IT} = P_A^I + \tau_A^I = (\alpha + \lambda_h + \mu_e \lambda_f) T_{Af} + \delta \left(\frac{N_A}{s} \right) + \tau_A^I \tag{D.2}$$

It is obvious to find that since P_A^T and P_A^{IT} only differs in the intercepts, a fixed road toll with nudging information and the generalized travel cost function of metro users is the same, a fixed road toll with nudging information can always reach the same number of commuters in each mode at equilibrium. And under nudging information, the uninternalized external cost is lower than without information when the number of road travelers is the same in the two cases. Hence, to maximize social welfare, the combination of nudging information and a fixed road toll can always reach a higher maximum social welfare than only a fixed road toll, which can be summarized as:

$$SW_T < SW_{IT} \tag{D.3}$$

This completes the proof.

Appendix E. Proof of Proposition 4

Similar to the proof of Proposition 1, since the metro serves as an imperfect substitution for driving, the nudging information will increase the crowding effects in the metro system. In some extreme cases where shifting road commuters to the metro system is sufficiently expensive, the nudging information can finally produce negative effects.

This completes the proof of condition (a) in Proposition 4.

To proof the condition (b) in Proposition 4, it is necessary to avoid the potential negative effects of nudging information that are addressed in condition (a). Therefore, the following analysis is restricted to the impact of nudging information on the road mode, achieved by setting $\eta = 0$, which excludes the possible negative effects related to the metro system.

Specifically, we use the case discussed in Section 4.2 where commuters are heterogenous in environmental preferences and car types to prove Proposition 4.

When $\mu_{eH}\lambda_{qH} > \mu_{eL}\lambda_{qL}$, an intended swap occurs where the more polluting cars travel during center of peak while the cleaner cars travel during shoulder of peak. Specifically, the polluting cars endure a relatively longer queuing time than the cleaner cars under nudging information. And the difference of emission cost between the two group are unbounded since the value of λ_{qH} can be significantly larger than λ_{qL} holding $\mu_{eH}\lambda_{qH} > \mu_{eL}\lambda_{qL}$.

In the case where all the commuters are inelastic, the change of emission cost under nudging information can be derived as:

$$\Delta EC = \frac{\delta}{2s}(N_{AH}^2\lambda_{qH}(\frac{1}{\alpha+\lambda_h+\lambda_{qH}\mu_{eH}} - \frac{1}{\alpha+\lambda_h\mu_h}) + N_{AL}^2\lambda_{qL}(\frac{1}{\alpha+\lambda_h+\lambda_{qL}\mu_{eL}} - \frac{1}{\alpha+\lambda_h\mu_h}) +$$

$$N_{AH}N_{AL}(\frac{2\lambda_{qH}}{\alpha+\lambda_h+\lambda_{qL}\mu_{eL}} - \frac{\lambda_{qH}+\lambda_{qL}}{\alpha+\lambda_h\mu_h})) \tag{E.1}$$

When all the commuters are nearly indifferent to the environment ($\mu_{eH} \to 0$, $\mu_{eL} \to 0$) and already know the exact health cost caused by driving without information ($\mu_h = 1$), ΔEC can be further simplified as:

$$\Delta EC = N_{AH}N_{AL}\left(\frac{\lambda_{qH}-\lambda_{qL}}{\alpha+\lambda_h}\right) > 0 \tag{E.2}$$

Hence, by changing the value of λ_{qH} and λ_{qL}, the extra emission cost brought by nudging information is unbounded in this case which can finally hurt the whole system.

This completes proof of the condition (b) in Proposition 4.

Appendix F.

Additional case I- $\frac{1}{\lambda_q}$ *and* $\frac{1}{\lambda_f}$ *vary in proportion* α

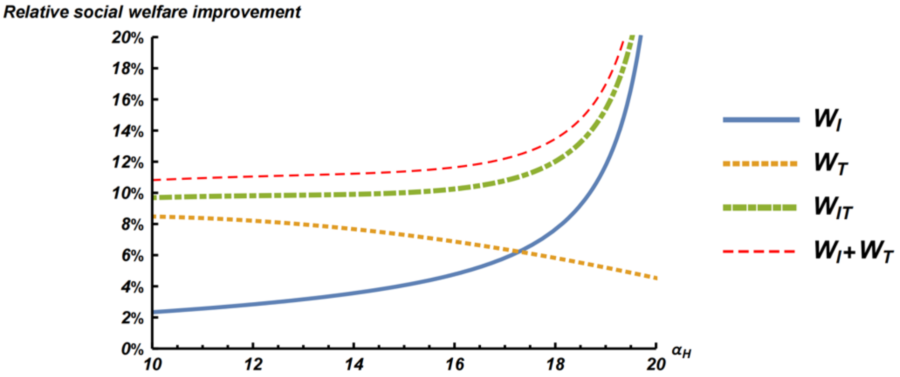


Figure F.1. Relative welfare improvement by information incentives under varying $\alpha_H(\frac{\alpha_L+\alpha_H}{2} = 10)$

Additional case II- ${\mu_e}^2$ *varies in proportion* α

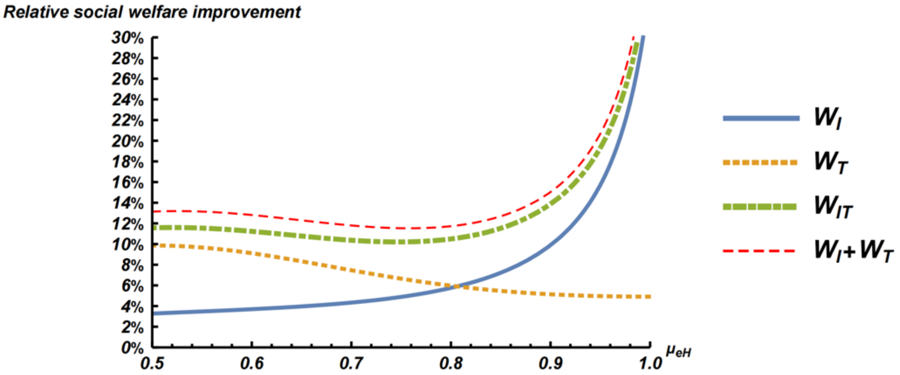


Figure F.2. Relative welfare improvement by information incentives under varying μ_{eH} $(\frac{\mu_{eH}+\mu_{eL}}{2} = 0.5)$

Additional case III- ${\mu_e}^{\frac{1}{2}}$ *varies in proportion* α

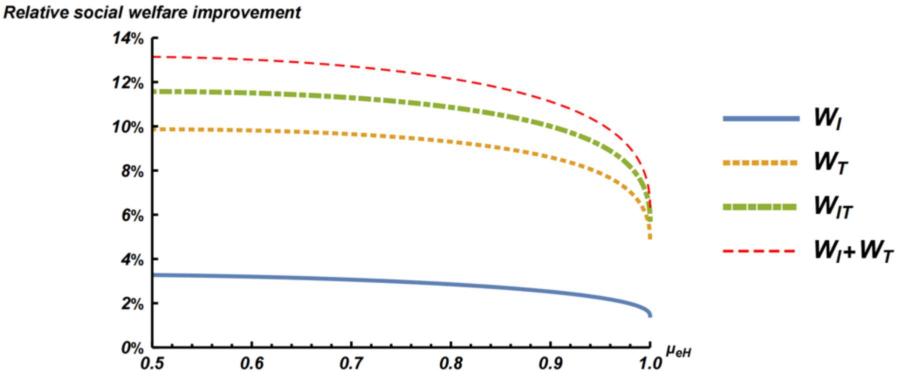
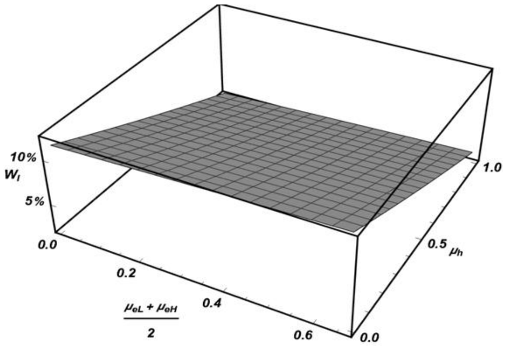
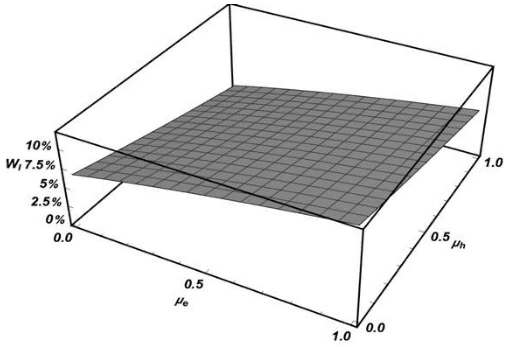


Figure F.3. Relative welfare improvement by information incentives under varying μ_{eH} ($\frac{\mu_{eH}+\mu_{eL}}{2}=0.5$)

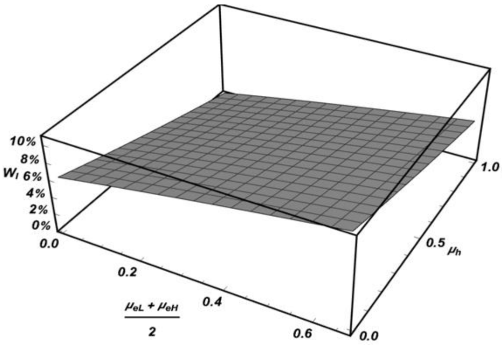
Appendix G.



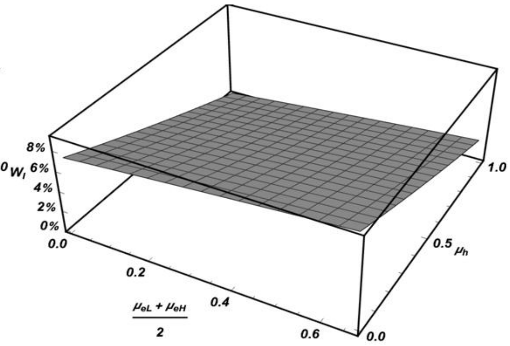
(a) λ_f and λ_q vary in proportion $(1-\mu_e)^2$



(b)$\frac{1}{\lambda_q}$ and $\frac{1}{\lambda_f}$ vary in proportion α

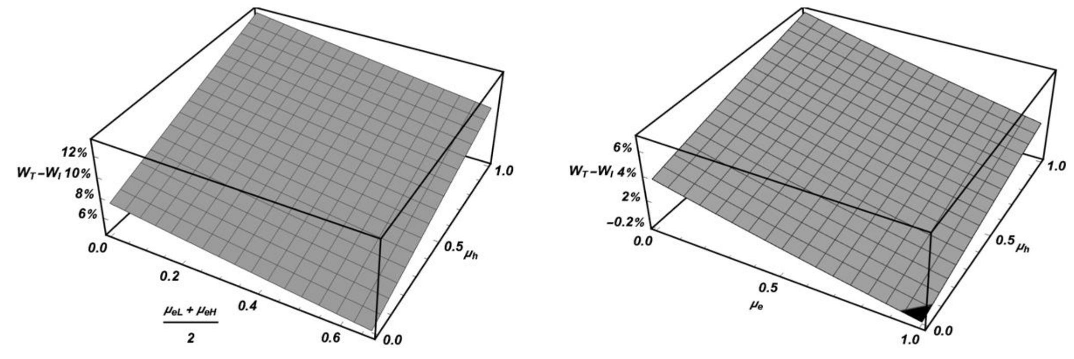


(c) $\mu_e{}^2$ varies in proportion α



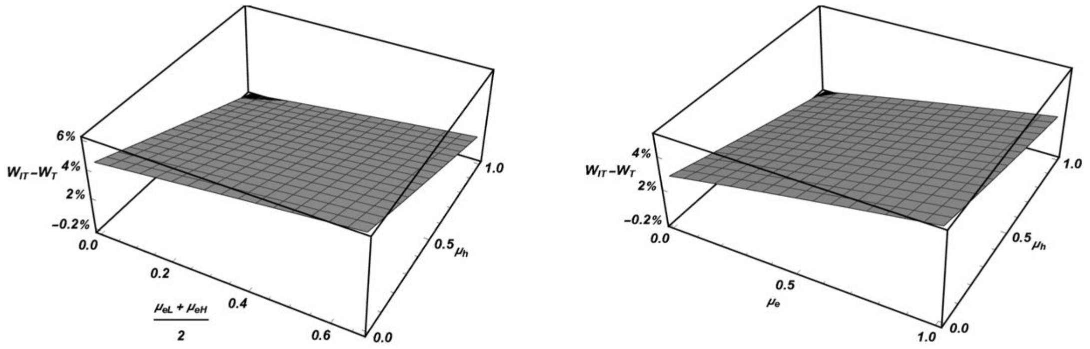
(d) $\mu_e{}^{\frac{1}{2}}$ varies in proportion α

Figure G.1. Relative welfare improvement by nudging information under varying μ_e, and $\mu_h(\mu_{eH} = 3\mu_{eL})$



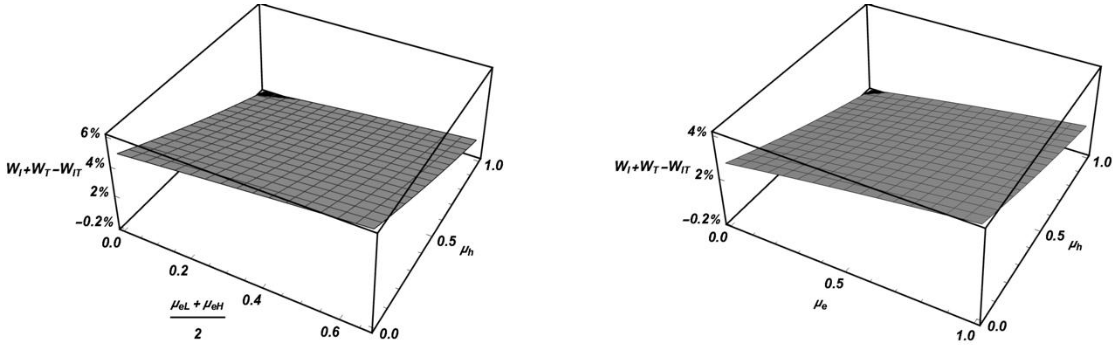
(a) λ_f and λ_q vary in proportion $(1-\mu_e)^{\frac{1}{2}}$ (b) ${\lambda_f}^2$ and ${\lambda_q}^2$ vary in proportion α

Figure G.2. Comparison between W_I and W_T under varying μ_e and $\mu_h(\mu_{eH} = 3\mu_{eL})$



(a) λ_f and λ_q vary in proportion $(1-\mu_e)^{\frac{1}{2}}$ (b) ${\lambda_f}^2$ and ${\lambda_q}^2$ vary in proportion α

Figure G.3. Comparison between W_{IT} and W_T under varying μ_e and $\mu_h(\mu_{eH} = 3\mu_{eL})$



(a) λ_f and λ_q vary in proportion $(1-\mu_e)^{\frac{1}{2}}$ (b) ${\lambda_f}^2$ and ${\lambda_q}^2$ vary in proportion α

Figure G.4. Comparison between $W_I + W_T$ and W_{IT} under varying μ_e and $\mu_h(\mu_{eH} = 3\mu_{eL})$