WORKING PAPER

ITLS-WP-23-16

A novel mobility consumption theory for road user charging

By
Michiel C.J. Bliemer[a], Allister Loder[b], Zuduo Zheng[c]

[a] Institute of Transport and Logistics Studies (ITLS),

The University of Sydney, Australia

[b] School of Engineering and Design, Technical University of Munich, Germany

[c] School of Civil Engineering, The University of Queensland, Brisbane, Australia

August 2023

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in

Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-23-16

TITLE: A novel mobility consumption theory for road user charging

ABSTRACT: Building on the analogy between electrical energy and mobility, we propose a novel mobility consumption theory based on the idea of the required reserved space headway of vehicles while driving. In this theory, mobility is ``produced'' by road infrastructure and is ``consumed'' by drivers in a similar fashion to power that is produced in power plants and consumed by electrical devices. The computation of mobility consumption only requires travel distance and travel time as input, as well as two physical parameters that are readily available, namely vehicle length and reaction time. We argue that mobility consumption is a more comprehensive measure for road use than travel distance (or travel time) alone as it captures road use over both space and time. One application area for our mobility consumption theory that we look at in this study is road user charging. We use mobility consumption theory to develop a mobility-based charging scheme as a novel road pricing approach and compare it to distance-based charging in two case studies. When considering only departure time choice in a simple bottleneck model, we show that mobility-based charging can reduce congestion akin a congestion pricing scheme, unlike distance-based charging. Further, when considering route choice, we show that distance-based charging can increase congestion as it encourages drivers to take shortcuts through routes with low capacity, while mobility-based charging mitigates this effect. The proposed mobility-based charging scheme is further capable of considering technological innovation in vehicle automation and carbon charging.

KEY WORDS: ***road use; mobility consumption; road pricing reform; road user charging***

AUTHORS: **Bliemer, Loder, Zheng**

ACKNOWLEDGEMENTS: N/A

CONTACT: INSTITUTE OF TRANSPORT AND LOGISTICS STUDIES (H04)

The Australian Key Centre in Transport and Logistics Management
The University of Sydney NSW 2006 Australia
Telephone: +612 9114 1813
E-mail: business.itlsinfo@sydney.edu.au
Internet: http://sydney.edu.au/business/itls

DATE: August 2023

1 Introduction

Road pricing, also referred to as road user charges, consist of all direct costs to road users. The primary objectives of such road user charges are to generate revenue for fund the construction of new infrastructure, to fund maintenance of roads and/or to subsidise public transport, but are increasingly considered for managing travel demand, alleviating congestion, and reducing emissions. Pricing in road transport has been widely studied by economists, engineers, and psychologists to assess the impact on travel behaviour, traffic flow, acceptability and equity (see e.g., Verhoef et al., 2008)

Transport economists typically argue in favour of first-best pricing, where road prices reflect the social marginal cost of a private transport vehicle (Small and Verhoef, 2007; Pigou, 1920; Lindsey and Santos, 2020). In practice, only second-best pricing is achievable, of which area-based and distance-based charging are prominent examples. Area-based (cordon) charging systems have been introduced in a select number of cities (e.g., Singapore, London, Stockholm) and aim to reduce congestion locally. Area-based charging is a form of congestion charging and an overview of congestion pricing methods and technologies is provided in de Palma and Lindsey (2011).

In this paper we focus on road pricing reform where existing road pricing schemes are replaced with a more direct way of charging road users, of which distance-based charging has received most attention (see e.g., Hensher and Mulley, 2014). Most countries use a combination of annual registration fees and fuel excise taxes to (partly) pay for the construction and maintenance of road infrastructure. Annual vehicle registration fees, also referred to as motor vehicle tax, are often considered unfair and inefficient because they do not depend on road use, while fuel excise taxes are financially unsustainable because revenues are decreasing due to the rise of electric vehicles. For these reasons, governments around the world are looking to replace existing road user charges with what is often referred to as a 'user pays' system. Distance-based charging systems have been introduced in several European countries (e.g., Germany, Belgium) for heavy vehicles. Distance-based charging for cars has been discussed for many years in several countries, but has only been introduced on a small scale in a few countries.

While distance-based charging is widely considered to provide a direct relationship with road use, in this paper we argue that total travel distance, also referred to as vehicle kilometer/miles travelled (VKT/VMT), is an incomplete measure for road use since it ignores the time dimension. It is well-known that analysing travel or traffic requires the consideration of both space and time (Greenshields, 1935; Daganzo, 2007). Therefore, in this study, we propose a novel theory where drivers are charged for their actual road use in space and time, considering that they claim, or 'consume', a certain amount of road space at each moment in time. This theory is generic and can be applied to all road types and current and future vehicle types. We show that this translates into charges for travel time and travel distance where the rates depend on the vehicle type, a type of road user charging that we will refer to as mobility-based charging.

We make an analogy between mobility systems and electrical energy systems and argue that mobility consumption can be measured in terms of kilometre-hours (kmh) analogous to electricity consumption in kilowatt-hours (kWh). Establishing road use measured in kmh is a key contribution of this paper and forms the basis of a novel road user charging scheme that we refer to as mobility-based pricing. Several analogies between traffic flow to physical approaches have been made in

the past, in particular applying physical laws of particles, fluids or gases to traffic flow to the flow (Helbing, 2001; van Wageningen-Kessels et al., 2015; Newell, 1993; Herman and Prigogine, 1979), but our analogy between electrical energy and mobility is novel.

In Section 2 we first present this analogy. In Section 3 we outline our novel mobility consumption theory in more detail and provide a mathematical derivation that decomposes mobility consumption into two components, where the first component depends on travel distance and the second component depends on travel time. In this section we also illustrate mobility consumption graphically in a space-time diagram. In Section 4 we apply our mobility consumption concept to develop a novel road user charging scheme in which drivers pay for their actual road use in space and time, which we refer to as mobility-based charging. In Section 5 we show the impact of this novel road user charging scheme on departure time and route choice in two simple case studies. In Section 6 we illustrate how vehicle automation and carbon emissions can be incorporated into mobility consumption and mobility-based charging. We end this analysis with a discussion and conclusions in Section 7.

2 An analogy between transport and electricity systems

We start by proposing an analogy that shows that power systems and transport systems have similar characteristics from a production and consumption perspective. However, it should be explicitly stated that we do not claim that these two systems are physically identical or are governed by identical mathematics.

First, we present an analogy between (electrical) energy and mobility in Section 2.1, followed by Sections 2.2 and 2.3 where we outline similarities between power consumption and production and mobility consumption and production. Finally, we briefly discuss the loss of energy and mobility in Section 2.4.

2.1 Energy versus mobility

Energy is defined as the ability to do work and in this paper we focus on electrical energy. End-users consume electrical energy to power their devices during a certain period of time. *Power* is measured in watt (W), which is equal to 1 Joule per second and describes the rate at which energy is consumed when running an electrical device. For example, a 60W lightbulb uses 60 Joules of energy every second when it is turned on. *Power consumption* is typically measured in kilowatt-hour (kWh) and billed by energy providers to end-users based on readings from electricity meters. Electricity meters typically have two readings, namely one for peak (day-time) power consumption and one for off-peak (night-time) power consumption, which are often charged at different rates.

Mobility is defined as the ability to move and in this paper we focus on mobility via motorised road transport. End-users consume mobility when they have a transport need. *Road space* is measured in meters (m) and describes the rate at which mobility is consumed when using roads. For example, a car driving on a motorway in free-flow uses around 30 m of claimed safe space headway, including the length of the vehicle, at each moment in time. *Mobility consumption* could therefore be measured in kilometer-hour (kmh) and billed by road operators to end-users.

Electricity system	Transport system
Electrical energy (kWh)	Mobility (kmh)
Power	Road space
Run electrical device	Transport need
Power plant	Road infrastructure
Power generator	Lane
Blackout	Congestion, Gridlock
Electron	Vehicle
Heat	Delay

Table 1: Comparison of power systems and transport systems

Odometers in vehicles capture distance travelled but do not measure road space use over time.

The respective matching of elements is summarised in Table 1 and will be further discussed in the following subsections.

2.2 Supply

Power stations convert primary sources such as crude oil, coal, and wind into electric energy. The *installed power capacity* is the maximum rate at which electricity can be generated by a power station and is typically expressed in kilowatt (kW) or megawatt (MW). *Power production* is measured in kWh, megawatt-hour (MWh), or gigawatt-hour (GWh), and describes how much electric energy is generated by a power station in one year. The ratio of actual power production and the amount of energy the power station would have produced at full capacity is called the power capacity factor. Baseload power generators are always on to provide a constant power supply but cannot be easily turned off or quickly vary their output, while peaking power generators can be used on-demand. For example, Eraring power station, the largest in Australia, has an installed power capacity of 2,922 MW. Therefore, it could have a power production of $2922 \cdot 365 \cdot 24 = 25,597$ GWh per year if run at full capacity. The actual power production of Eraring power station is on average 16,012 GWh, which means a power capacity factor of 63%. It has four 720 MW coal-fired baseload generators and one 42 MW diesel peaking generator.

Road infrastructure converts primary sources such as land into mobility. The *installed mobility capacity* can be defined as the maximum rate at which road space can be generated by a piece of road infrastructure and can be expressed in kilometre (km). *Mobility production* can therefore be measured in kilometre-hour (kmh) and describes how much mobility is generated by road infrastructure in one day. The ratio of actual mobility production and the amount of mobility road infrastructure would have produced at full capacity can be called the mobility capacity factor. Road infrastructure can have multiple *mobility generators*, typically referred to as lanes. Baseload mobility generators are lanes that are always open for driving in their default direction, while peaking mobility generators can be used on-demand, e.g., reversible lanes or shoulder lanes that are open during peak hours. For example, the Sydney Harbour Bridge (northbound) is 1.15 km in length and has at most 5 lanes available in this direction, therefore it has an installed mobility capacity of

5.75 km and can therefore have a mobility production of $5.75 \cdot 24 = 138$ kmh per day at full capacity. The actual mobility production of the Sydney Harbour bridge (towards north) varies between 40% and 100% depending on lane availability throughout the day. It has 2 baseload lanes that are permanently open northbound and 3 reversible peaking lanes that can be switched to northbound to accommodate tidal flow.

2.3 Demand

Power consumption varies over the day and from day to day. On a typical work day, the power consumption rate peaks during the early evening. Expected power consumption influences how much power is produced. If the electricity demand is larger than the electricity supply then *blackouts* occur. Blackouts can be avoided through demand management measures. Loadshedding is a hard measure where access to electricity is cut for certain end-users, while information and encouragement to reduce power consumption are soft measures.

Mobility consumption also varies over the day and from day to day. On a typical work day, the mobility consumption rate typically peaks in the morning and evenings following commuting patterns. In most cases, the number of available lanes for driving in a certain direction on a road is fixed in the short term and therefore expected mobility consumption only influences mobility production in the long term. If the demand for mobility is larger than the mobility supply then *congestion* or even *gridlock* occurs. Congestion or gridlock can be avoided through demand management measures. Lane closures and ramp metering are examples of hard measures where access to mobility is cut or reduced for certain end-users, while information and encouragement to reduce car use are soft measures.

Production and consumption during a typical work day is graphically illustrated in Figure 1, where Figure 1(a) illustrate the rates at which power is produced and consumed and Figure 1(b) illustrates the rates for mobility. Power production generally follows power consumption where possible. In contrast, mobility production is fixed for most roads and is only fully consumed at locations where congestion occurs, in which case drivers exhibit car-following behaviour with safe space headways that cover the entire lane. Mobility consumption is discussed in more detail in Section 3.

2.4 Energy and mobility loss

Produced power is lost when not consumed, unless the energy is stored (e.g., in batteries or pumped hydro). Electrons in power lines are *carriers of electrical energy* and allow electric signals to travel almost at the speed of light. There could be *power transmission loss* due to heat in conductors or transformers. There could also be *power inefficiency loss*, namely not all consumed energy is used for the purpose of the device. For example, a typical lightbulb is only able to use a small amount of electrical energy to create light while the majority is released as heat.

Produced mobility is also lost when not consumed, but in contrast to power, mobility cannot be stored. Vehicles in transport networks are *carriers of mobility* and allow people and goods to travel at maximum vehicle/road speeds. There could be *mobility transmission loss* due to persistent or non-persistent delays at bottlenecks or intersections, or due to driving behaviour (e.g., not using

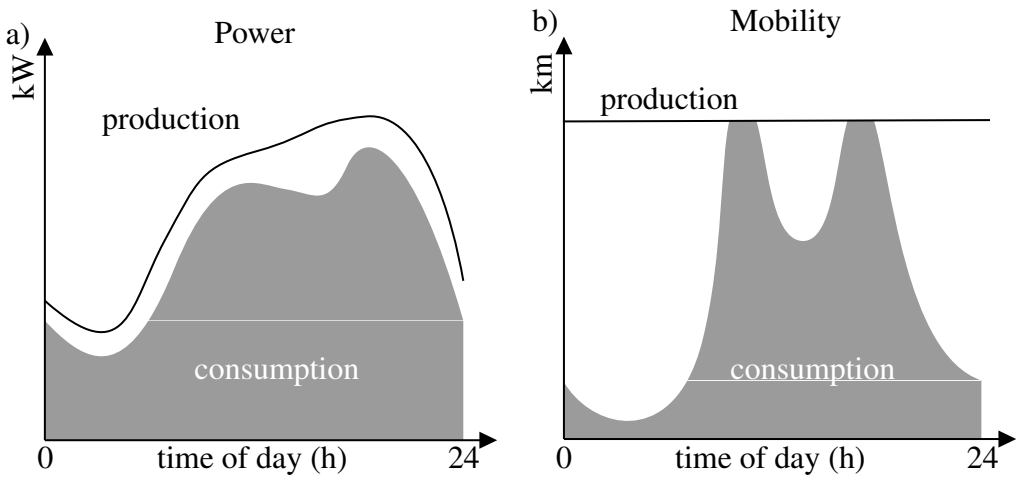

Figure 1: Graphical illustration of power and mobility production and consumption rates.

available road space in all lanes). There could also be *mobility inefficiency loss*, namely not all consumed mobility is used to satisfy transport needs. For example, a taxi driving around without passengers results in mobility loss due to dead running.

3 A novel mobility consumption theory

In this section we present our novel mobility consumption theory to describe road use by making a direct relationship with safe space headway. We first derive this relationship mathematically and show that mobility consumption depends on travel distance and travel time using vehicle-specific parameters. Next we illustrate mobility production and consumption graphically using space-time diagrams, and finally we describe how to obtain the parameters for computing mobility consumption.

In Section 3.1 we define mobility consumption by relating road use to space headway that drivers consume over time while driving, and we show that mobility consumption can be computed based on travel time and travel distance. In Section 3.2 we visualise mobility production and consumption in space-time diagrams using empirical data using vehicle-specific travel time and distance factors estimated in Section 3.3.

3.1 Safe space headway and mobility consumption

Consider a trip of vehicle n and let $x_n(t)$ describe its space-time trajectory, where vehicle n starts the trip at location x_n^0 at time instant t_n^0, i.e., $x_n\left(t_n^0\right) = x_n^0$, and finishes at location $x_n^0 + D_n$ at time $t_n^0 + T_n$, i.e., $x_n\left(t_n^0 + T_n\right) = x_n^0 + D_n$, where D_n is the travel distance and T_n is the travel time. Let $h_n(t)$ be the safe space headway of vehicle n at time instant t, which is the amount of reserved road space for safety reasons. Table 2 summarises the variables and parameters used in Sections 3 and 4.

Safe space headway is an increasing function of vehicle speed $v_n(t) = dx_n(t)/dt$ since driving

Variable/parameter	Unit	Description
T_n	h	Travel time of vehicle n
T_n^0	h	Minimum (free-flow) travel time of vehicle n
D_n	km	Travel distance of vehicle n
M_n	kmh	Mobility consumption of vehicle n
$x_n(t)$	km	Location of vehicle n at time instant t
$v_n(t)$	km/h	Speed of vehicle n at time instant t
$h_n(t)$	km	Safe space headway for vehicle n at time instant t
τ_n	h	Reaction time of (the driver of) vehicle n
λ_n	km	Minimum space headway for vehicle n
c_n	\$/kmh	Unit charge for mobility consumption
μ_n	\$/h	Travel time charging rate
δ_n	\$/km	Travel distance charging rate

Table 2: List of variables and parameters

at high speeds requires a larger gap to the vehicle in front. Assuming a linearly increasing function of vehicle speed results in the safe space headway function proposed by Pipes (1953),

$$h_n(t) = \lambda_n + \tau_n v_n(t), \tag{1}$$

where $\lambda_n > 0$ is the minimum space headway for vehicle n, which equals the length of the vehicle plus the average space between the rear bumper of the front vehicle and the front bumper of the subject vehicle when standing still, and where $\tau_n > 0$ is the reaction time needed for vehicle n to adjust its speed against a stimulus without deliberate delay (Sharma et al., 2109). This simple linear relationship between speed and spacing is a commonly used assumption in traffic flow theory (e.g. it is used in Newell's (2002) simplified car following model), consistent with the triangular fundamental diagram and also supported by empirical evidence (Ma and Ahn, 2008). Although more sophisticated nonlinear relationships have been estimated, a linear relationship allows us to derive a practically useful expression for (an approximation of) mobility consumption.

Figure 2 illustrates Eqn. (1), where traffic conditions B, C, and D are in the congested branch such that all vehicles are car-following and distances between vehicles are determined by the safe space headway. Traffic condition A is in free-flow in which vehicles are not car-following with a space headway larger than the safe headway.

The reserved road space of a vehicle of vehicle n along their trajectory, which we refer to as mobility consumption, can be derived as

$$M_n = \int_{t_n^0}^{t_n^0+T_n} h_n(t)\,dt = \int_{t_n^0}^{t_n^0+T_n} \lambda_n dt + \int_{t_n^0}^{t_n^0+T_n} \tau_n v_n(t)\,dt = \lambda_n T_n + \tau_n x_n(t)\Big|_{t_n^0}^{t_n^0+T_n} = \lambda_n T_n + \tau_n D_n. \tag{2}$$

Mobility consumption M_n therefore directly depends on the total travel time and travel distance, where λ_n and τ_n are vehicle-specific and physically meaningful input parameters. Note that travel time T_n [h] is multiplied by λ_n [km] and travel distance D_n [km] is multiplied by τ_n [h],

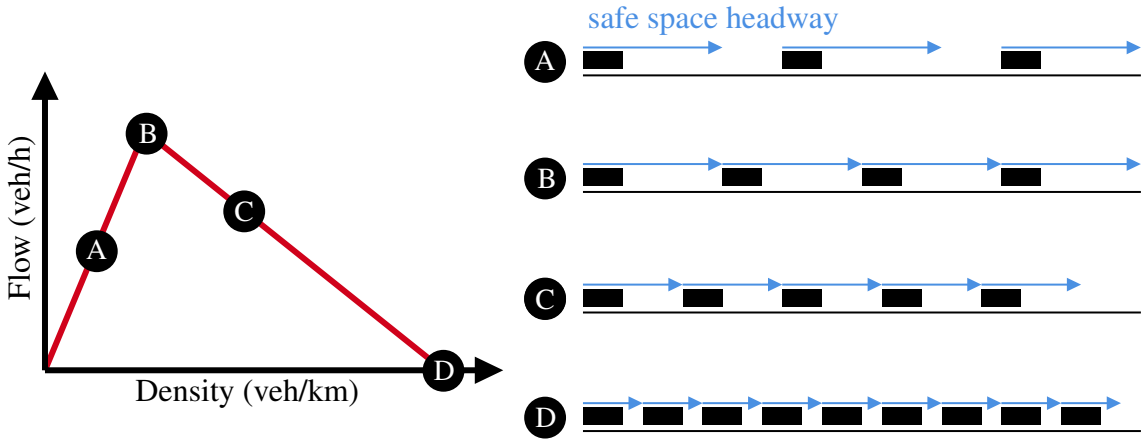

Figure 2: Reserved road space dependent on traffic conditions

hence mobility consumption M_n is measured in kmh. The total mobility consumption in a transport system is simply $\sum_n M_n$, which can be calculated based on the total travel time and total travel distance of all vehicles.

3.2 Mobility consumption in space-time diagrams

Instead of looking at mobility consumption of a single vehicle, it is also possible to determine mobility consumption for a specific area Δx and a specific time period Δt by considering only trajectories in this specific space-time domain. Although this is outside the scope of this paper, such an analysis of mobility production and consumption for a specific road segment at a certain time of day may be useful for determining road use efficiency and traffic management.

Mobility consumption for a given road segment and time period is graphically illustrated using space-time diagrams of a road with a single lane in Figure 3. Trajectories of vehicles with a constant speed are shown with a solid black line while the grey areas indicate their mobility consumption, i.e., their safe space headway. In free-flow situations as shown in Figure 3(a) vehicles are spaced out and the reserved road spaces of vehicles are not touching each other. This situation changes when the flow is at capacity and the reserved road spaces of vehicles become adjacent as seen in Figure 3(b). Once congestion emerges, vehicles slow down and a queue grows along a backward shockwave, see Figure 3(c). When vehicles drive slower in a queue, their reserved road space per time unit decreases as per Eqn. (1), but since they spend more time travelling the same distance their overall mobility consumption will actually increase. This additional mobility consumption is shown in Figure 3(d), where area "I" corresponds to the minimum mobility consumption and area "II" corresponds to the additional mobility consumption due to congestion. The latter can be seen as "transmission losses" in the transport system. For the considered road and time period, the mobility consumption for each vehicle n can therefore be decomposed into a minimum component, M_n^{I}, and an additional component, M_n^{II}, where M_n^{I} is based on free-flow

travel time T_n^0,

$$M_n = M_n^{\mathrm{I}} + M_n^{\mathrm{II}}, \qquad \text{where} \quad M_n^{\mathrm{I}} = \lambda_n T_n^0 + \tau_n D_n \quad \text{and} \quad M_n^{\mathrm{II}} = \lambda_n \left(T_n - T_n^0\right). \tag{3}$$

Further, Figure 3(e) illustrates the associated mobility production and consumption rates over time, where the mobility production (i.e., road space) is fully consumed during a short period of time.

To illustrate mobility production and consumption in practice, we analysed the Next Generation Simulation (NGSIM) I-80 data set (U.S. Department of Transportation Federal Highway Administration, 2016) and plotted trajectories in space-time diagrams, see Figure 4. The first exhibit (Figure 4(a)) shows trajectories on the first lane with almost free-flow conditions, while the second, while the second exhibit (Figure 4(b)) shows trajectories on the fifth lane in saturated and congested conditions. Each exhibit shows a road space of 300 meters, i.e., $\Delta x = 0.3$ km, and a time duration of three minutes, i.e., $\Delta t = 0.05$ h, therefore in both depicted exhibits the mobility production is $\Delta x \cdot \Delta t = 0.015$ kmh.

The associated mobility consumption is shown in black in Figure 5 using values for λ_n and τ_n reported in Section 3.3). It can be clearly seen that in free-flow conditions on lane 1 a lot of mobility production has not been consumed (the remaining white space), while in saturated or congestion traffic situations one lane 5 most of the mobility production has been consumed.

3.3 Travel time and distance factors

Our proposed mobility consumption theory is parsimonious with only two vehicle-specific parameters, namely λ_n as the factor for travel time and τ_n as the factors for travel distance. In this section, we show how to estimate these two parameters using vehicle trajectory data. For illustration purposes, we look again at the NGSIM I-80 data set. Ma and Ahn (2008) confirmed the existence of a strong overall linear relationship between spacing and speed for passenger cars in this data set. Building on their analysis, we further analyse the distributions of vehicle length, bumper-to-bumper spacing (gap) and time headway as shown in Figure 6 to derive implications for our mobility consumption theory.

The distribution of vehicle length, which is part of λ_n, is depicted in Figure 6(a). This vehicle length distribution shows that the majority of vehicles is shorter than 5 m, with an average length of 4.87 m, suggesting that traffic comprises mostly of passenger cars. Together with an average bumper-to-bumper spacing of 6.25 m as reported by Ma and Ahn (2008), this yields an average space headway of 11.12 m, or an average of $\lambda_n = 11.12/1000$. Intuitively, we find in Figure 6(b) an increase in the average bumper-to-bumper spacing with increasing speeds. At speeds of less than 10 km/h, vehicles have an average bumper-to-bumper spacing of 5.76 m. The difference to the 6.25 m reported by Ma and Ahn (2008) is due to the fact that the 6.25 m is an estimated intercept of a linear model using all data, while a spacing of 5.76 m is derived from the empirical distribution at speeds lower than 10 km/h. Consequently, for the I-80 section we can derive an average of $\lambda_n = 10.63/1000$. It should be noted that some administrations (e.g., Queensland Government, 2022) require heavy vehicle drivers to drive under some circumstances at a safe following distances substantially larger than the general car-following values, inevitably increasing mobility consumption.

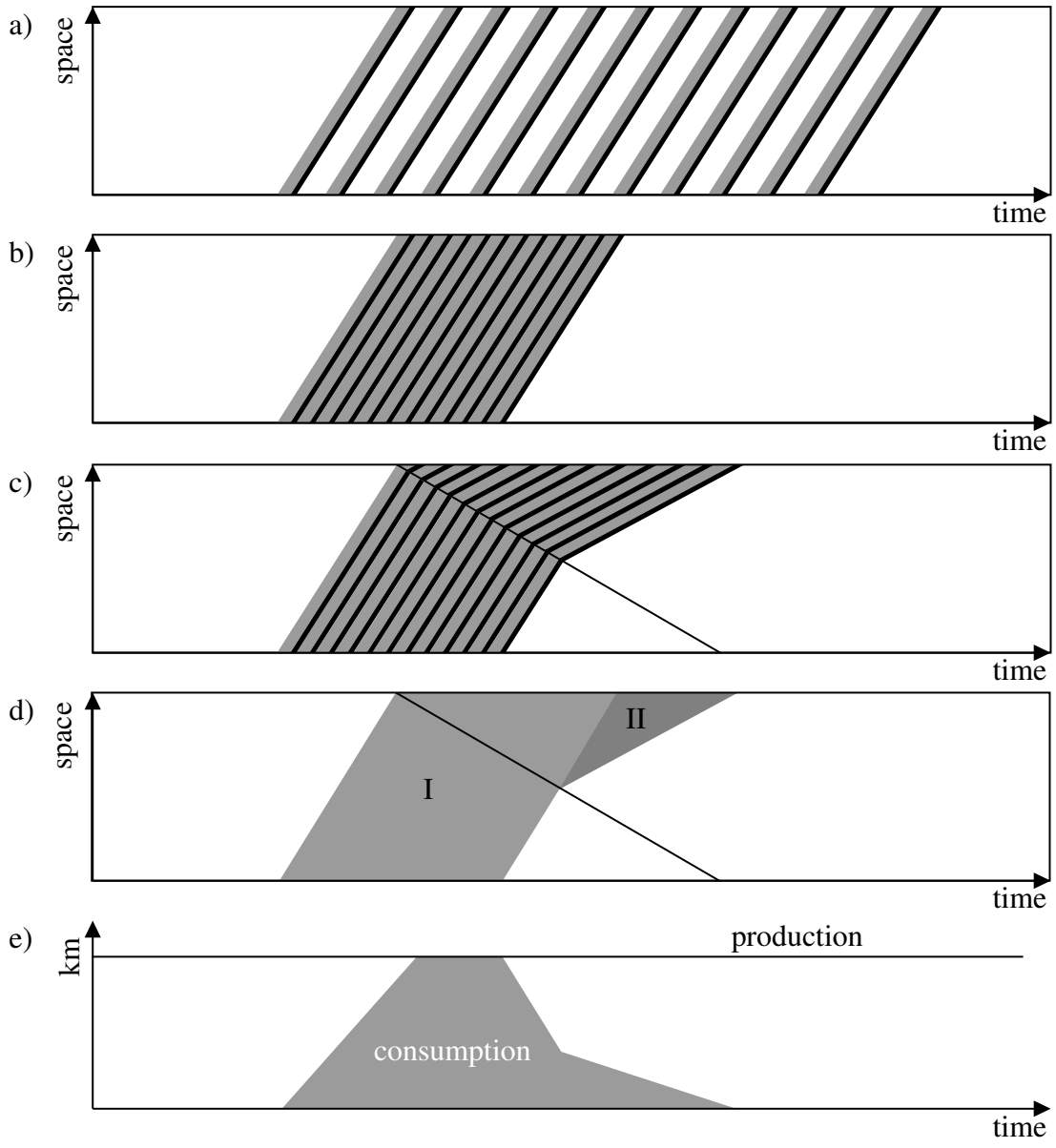

Figure 3: Reserved road space (a) in free-flow, (b) at flow capacity, (c) in congestion, (d) at minimum (area I) and additional (area II); (e) mobility production and consumption rate over time.

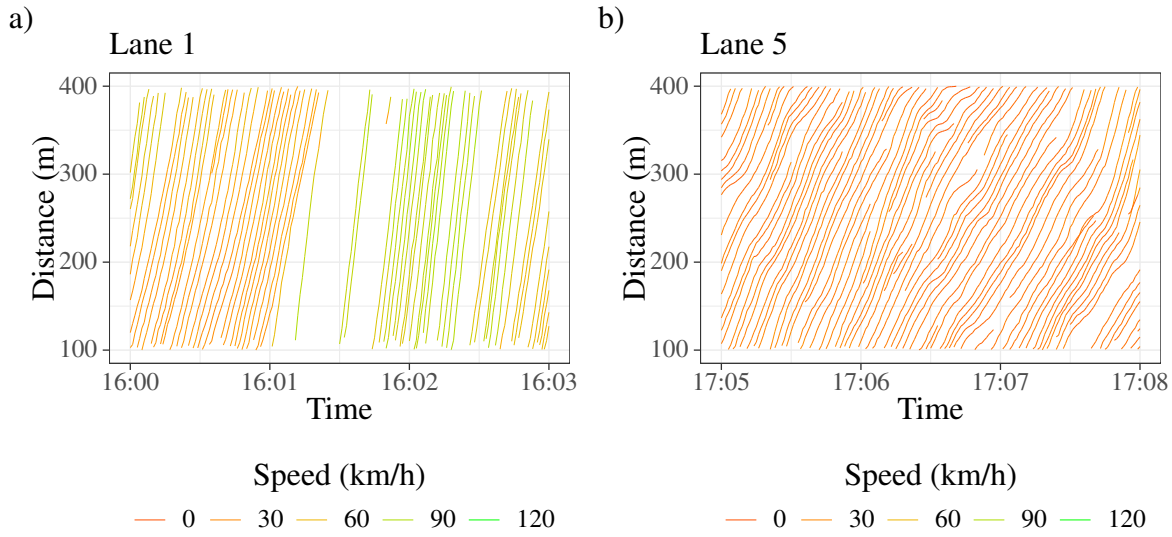

Figure 4: Trajectories from the NGSIM Interstate 80 data (U.S. Department of Transportation Federal Highway Administration, 2016).

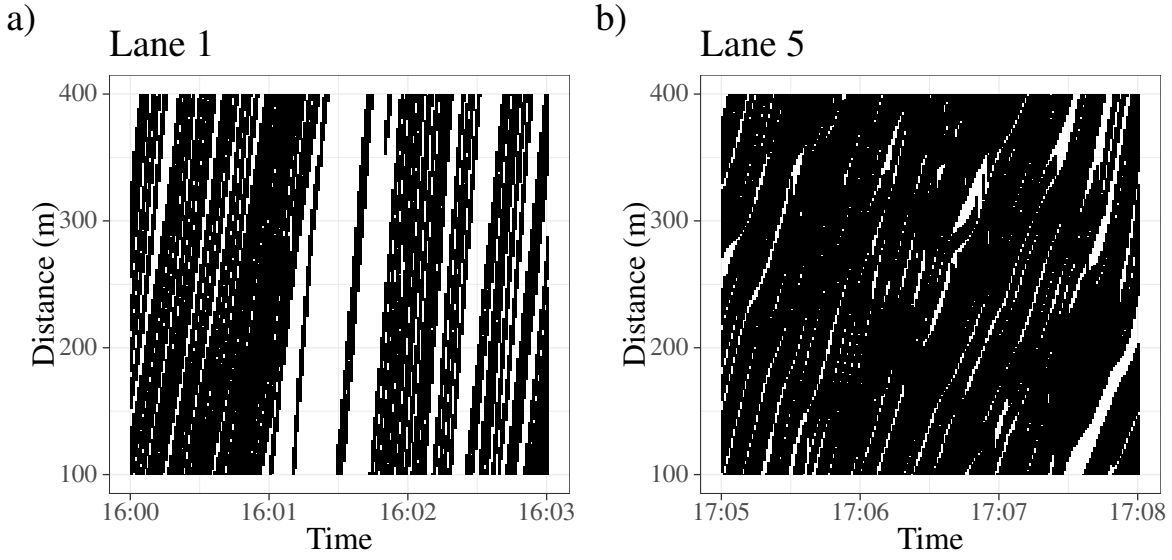

Figure 5: Mobility consumption based on trajectories from the NGSIM Interstate 80 data (U.S. Department of Transportation Federal Highway Administration, 2016). Black coloring indicates that the respective space-time cell has been consumed for mobility.

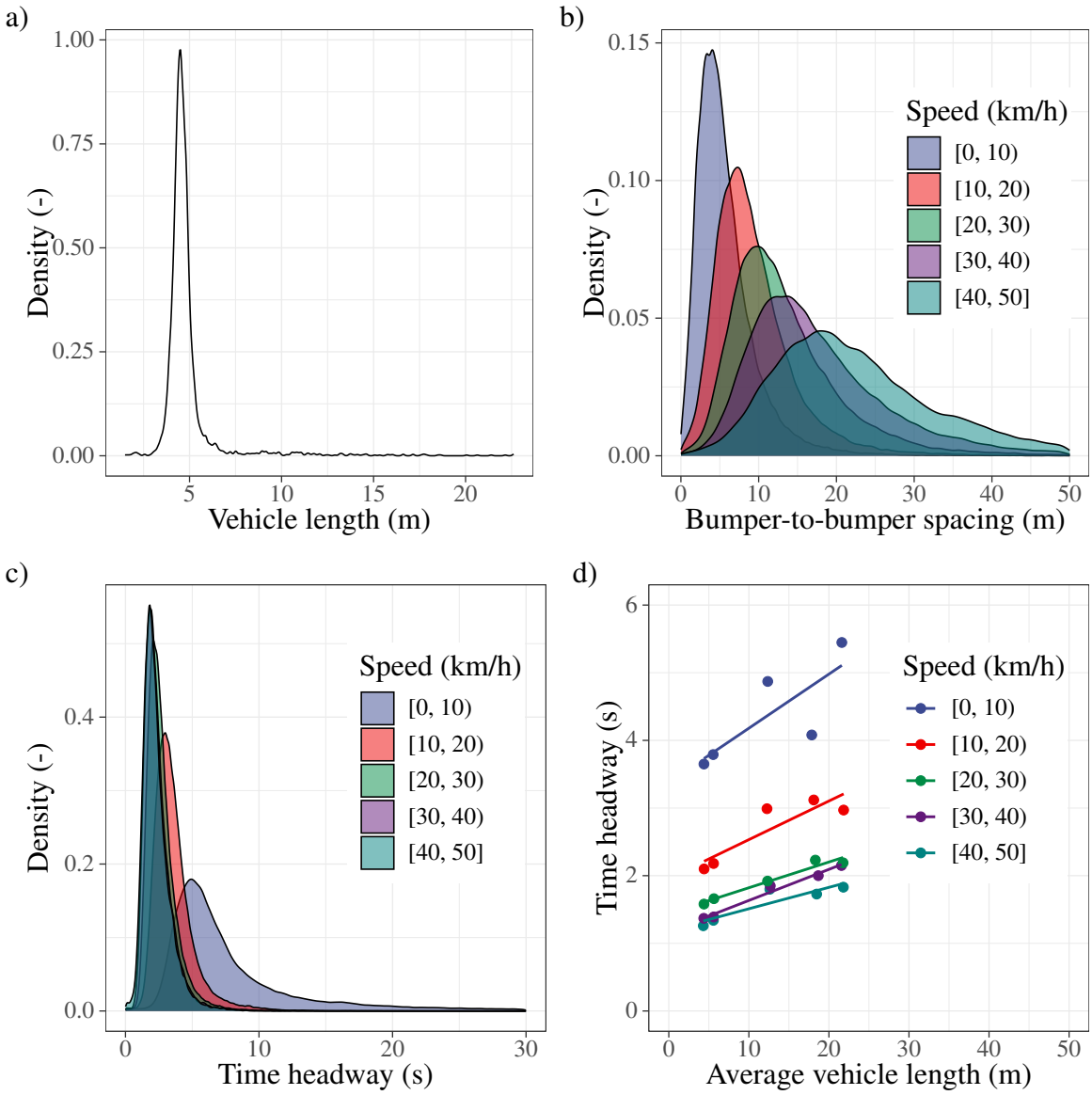

Figure 6: Distribution of variables derived from the Interstate 80 section in NGSIM data that are relevant for determining values of λ_n and τ_n.

Parameter	Unit	Value
Average bumper-to-bumper spacing, taken from Ma and Ahn (2008)	m	6.25
Average bumper-to-bumper spacing, derived from the data	m	5.76
Average vehicle size, derived from the data	m	4.87
Average reaction time, taken from Ma and Ahn (2008)	s	1.23

Table 3: Average values of spacing, vehicle size, and reaction time based on NGSIM I-80 data.

In the traffic flow literature, reaction time is often estimated through correlation analysis for a pair of vehicles built on the fundamental stimulus–response relationship either in the time domain (Zheng et al., 2013) or in the wavelet (Zheng et al., 2011a,b) or frequency domain (Li et al., 2010). However, for the purpose of our novel mobility consumption theory, obtaining accurate reaction time for each driver is not critical, as this theory is meant to be applied more generically. Thus, it is reasonable to use minimum time headway to approximate reaction time, which benefits from the fact that obtaining minimum time headway from vehicle trajectories is relatively straightforward. Observe in Figure 6(c) that time headway is increasing with decreasing speed and that headway shows substantial variation. Thus, to better understand the minimum time headway or τ_n we derive in Figure 6(d) the 10th percentile of the time headway distribution as a function of speed and vehicle size. Here, it can be seen that the minimum time headway approaches 1.23 as reported by Ma and Ahn (2008), i.e., $\tau_n = 1.23/3600$, with increasing speeds and decreasing vehicle length. Nevertheless, data suggests that τ_n increases with vehicle size.

A summary of the values of underlying factors λ_n and τ_n based on I-80 data is shown in Table 3. In general, When applying our mobility consumption theory, typical values reported in the literature can be used. For passenger cars this means that λ_n is between 0.010 (10 m) and 0.012 (12 m) and τ_n is between $1/3600$ (1 s) and $2/3600$ (2 s) (Green, 2000; Ranjitkar et al., 2003; Zheng et al., 2013).

4 Application to road user charging

In Section 4.1 we first briefly describe existing road pricing schemes and some of their challenges. Then we propose in Section 4.2 a novel charging scheme based on mobility consumption as an extension of distance-based charging schemes, list advantages of such mobility-based charging schemes, and provide an illustrative example with numbers.

4.1 Existing road pricing schemes

Existing road user charges include annual registration fees, annual motor vehicle tax, fuel excise tax, road tolls, etc. These charges are often considered unsustainable, unfair, and economically inefficient.

Unsustainable because revenues from fuel excise tax will decline due to cars becoming increasingly electric (Konstantinou et al., 2022). Further, given that running costs of electric vehicles

are very low, drivers of electric vehicles are expected to travel more and longer distances (Hensher, 2020). Travel demand on road networks can be better managed if there is a direct relationship between road use and road user charges, which will become increasingly important when automated vehicles arrive in the future in order to avoid significant dead running (i.e., driving around empty).

Unfair because most countries charge a fixed annual registration fee and motor vehicle tax, independent of number of kilometres driven. Also, fuel excise taxes disadvantage lower income groups more since they generally travel longer distances (Steinsland et al., 2018) and they may not be able to afford fuel-efficient, hybrid, or electric vehicles. Further, some car drivers cannot easily avoid toll roads while others are surrounded by untolled motorways.

Economically inefficient because Vickrey, Nobel Laureate and 'father of congestion pricing', recommended that prices should reflect road use and depend on traffic conditions (Vickrey, 1993). Registration fees and motor vehicle tax do not charge for road use but rather for car ownership. Fuel excise tax does not provide a monetary incentive to avoid peak hours to allow more efficient use of scarce road infrastructure. While toll roads may be more efficient and some tolls vary by time of day, tolls only exist on a relatively small number of roads while most roads are toll-free.

To overcome some limitations of existing road pricing schemes, Singapore, London, Stockholm, and Milan have introduced cordon charging schemes where drivers pay to enter the city centre, varied by time of day, with the aim to locally alleviate congestion (Metz, 2018). *Distance-based charging* is generally considered a natural and economically efficient way of paying for use of all roads, and many governments around the world are considering such a scheme. Several countries in Europe (e.g., Germany and Belgium) have implemented country-wide distance-based charging schemes for trucks using on-board GPS units, although other countries (e.g., France) stopped its implementation due to challenges (Rigot-Müller, 2018). New Zealand has introduced distance-based charging for light diesel vehicles only, while the state of Oregon (USA) offers only drivers of electric and fuel-efficient vehicles the option to pay per mile instead of a fixed vehicle licensing and registration fee, both using available odometers in the vehicle to determine distance travelled.

While distance sounds like a good proxy for road use, we argued in Section 3 that it is an incomplete measure for road use since it lacks the time dimension. Charging only for distance would encourage drivers making shortcuts through a city centre or residential area instead of choosing a longer but often faster motorway or main road around it. Further, distance-based charging does not encourage drivers to avoid driving during peak hours unless kilometre rates are time varying and an on-board unit with GPS is installed to measure distances travelled at different rates. Instead, we propose to measure and pay for road use based on mobility consumption as introduced in this paper, which we refer to as *mobility-based charging* and depends on both travel distance and travel time and directly reflects road use in space and time.

4.2 Mobility-based charging

Similar to charging for power consumption per kWh, we propose to charge for mobility consumption per kmh. Let c_n denote the vehicle-specific unit charge for mobility consumption. Then

mobility-based charging means that each vehicle n pays a road charge equal to

$$c_n M_n = \mu_n T_n + \delta_n D_n, \tag{4}$$

where $\mu_n = c_n \lambda_n$ is the travel time charging rate in \$/h and $\delta_n = c_n \tau_n$ is the travel distance charging rate in \$/km. Note that *distance-based charging* is a special case of mobility-based charging where $\delta_n > 0$ and $\mu_n = 0$, while *travel time-based charging*[1] is also a special case of mobility-based charging where $\mu_n > 0$ and $\delta_n = 0$. In principle any rates could be selected for μ_n and δ_n, but for road user charging to be consistent with our proposed mobility consumption theory it needs to hold that $\delta_n > 0$ and $\mu_n = (\lambda_n/\tau_n)\delta_n$.

There are several advantages of mobility-based charging: (i) it provides a direct relationship to road use in terms of space and time since mobility consumption theory maps distance and travel time onto a single scale based on traffic flow theory, (ii) it is easily measurable each month or year based on total travel distance (odometer reading) and total travel time (i.e., total vehicle running time via an internal clock), (iii) it automatically makes driving during congested periods more expensive, (iv) it offers a balance between shorter and faster routes, (v) it allows intuitive setting rates for different vehicle classes with respect to length, weight, engine type, automation level, etc.

Motor vehicle tax or registration fees depend on the weight of the vehicle while electric vehicles may receive a discount, and fuel excise taxes depend (indirectly) on the weight of the vehicle and fuel efficiency of the engine. In the same way, unit charge c_n may also depend on similar vehicle characteristics. In this paper we primarily focus on charges for road use, but additional charges could be imposed to internalise external costs that are related to travel distance or travel time, see Section 6.2.

One could argue that charging for both distance and time as in Eqn. (4) is not new since this is how metered taxis operate in many countries. For taxis, the time rate μ_n mainly depends on labour costs, while distance rate δ_n mainly depends on vehicle operating costs. However, taxi costs are unrelated to road use and hence taxi rates cannot be used as the basis for road user charging. Also car sharing companies often charge for both distance and time. For example, GoGet in Australia charges an hourly rate as well as a kilometre rate for car use. Again, charging for car use is not the same as charging for road infrastructure use, so these rates cannot be used as the basis for road user charging, but it does illustrate that many travellers are already familiar with being charged for both distance and time.

As an illustrative example, consider a car driver in Sydney, Australia. A typical car driver pays around A\$1,200 per year for motor-vehicle tax, registration fees, and fuel excise taxes, and drives 13,700 km per year (pre-COVID) (Australian Bureau of Statistics, 2019). The average driving speed in Sydney is 59.6 km/h (Australian Automobile Association, 2019), so the average hours spent driving on the network is about 230 hours per year. This means that an average car driver has a mobility consumption of $11.12/1000 \text{ km} \cdot 230 \text{ h} + 1.23/3600 \text{ h} \cdot 13700 \text{ km} = 2.56 + 4.68 = 7.24$ kmh, based on values for λ_n and τ_n derived in Section 3.3. Note that travel distance contributes more to mobility consumption than travel time.

To make road pricing reform acceptable to both government and car drivers, Hensher and Bliemer (2014) suggest a road user charging scheme that is revenue neutral and starts as a voluntary

[1] The term 'time-based charging' often refers to time-varying charges, so here we explicitly mean charging for travel time.

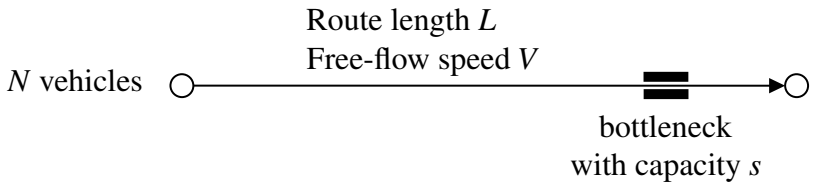

Figure 7: Single road with a bottleneck.

scheme where car drivers can receive a discount on their fixed registration fees and motor vehicle taxes when driving less. If the above-mentioned road user charges of A$1,200 per year are to be replaced by a revenue-neutral mobility-based charge, then car drivers will need to be charged an average of $c_n = 1200/7.24 = $ A$166/kmh. This corresponds to an average travel distance charging rate of $\delta_n = $ A$0.0567/km (5.67 cents per kilometre) and an average travel time charging rate of $\mu_n = $ A$1.846/h (3.08 cents per minute) using average parameter values for λ_n and τ_n derived in Section 3.3.

5 Impact of mobility-based charging on travel behaviour and congestion

In this section, we apply our mobility consumption theory to road user charging using two simple case studies. First, departure time choice in Section 5.1 and route choice in Section 5.2. These case studies illustrate how charging for mobility consumption influences behavioral choices and traffic outcomes. These case studies also compare outcomes of mobility-based charging, where $\delta_n > 0$ and $\mu_n = (\lambda_n/\tau_n)\delta_n$, with distance-based charging, where $\delta_n > 0$ and $\mu_n = 0$, and no charging, where $\delta_n = \mu_n = 0$.

In the following we assume N homogeneous travellers, each driving a vehicle of the same type such that we drop index n. We make the assumption that road user charging is introduced in a revenue neutral and hence cost neutral fashion and that demand is inelastic. Further, we consider traffic dynamics whereby mobility consumption and travel time are functions of departure time instant t.

In addition to the variables and parameters used in the previous sections, see Table 2, we use other variables and parameters summarised in Table 4.

5.1 Impact on departure time choice

Consider a single route of length L with a bottleneck at the end as shown in Figure 7. Since we are considering only a single route, we omit index i here from all variables. Further, consider N travellers that needs to decide when to drive their vehicle on this road taking into travel times and deviations from their preferred arrival time t^*. The well-known bottleneck model introduced by Vickrey (1969) and extended in Small (1982) and Arnott et al. (1993) describes the basic dynamics of road congestion and provides a behavioural description of departure time choice based on travel time delays and schedule delays. Schedule delays occur when arrival times deviate from a given

Variable/Parameter	Unit	Description
$C_i(t)$	\$	Generalised travel cost of route i when departing at time t
C_i^0	\$	Minimum generalised travel cost of route i
$r_i(t)$	veh/h	Departure flow rate on route i
$Q_i(t)$	veh	Vehicles encountered in queue on route i when departing at time t
s_i	veh/h	Bottleneck exit capacity on route i
L_i	km	Length of route i
V_i	km/h	Free-flow speed on route i
N	veh	Number of travellers/vehicles
t'	h	Departure time instant of the first traveller/vehicle
t''	h	Departure time instant of the last traveller/vehicle
t^*	h	Preferred arrival time
α	\$/h	Value of travel time
β	\$/h	Value of schedule delay arriving early
γ	\$/h	Value of schedule delay arriving late

Table 4: List of additional variables and parameters in case studies

preferred arrival time t^*. In this model, it is assumed that the private travel cost of a vehicle depends on its departure time and is composed of time-dependent travel time, schedule delay early, schedule delay late, and toll. Let $C(t)$ denote the generalised total travel cost, which can be defined as

$$C(t) = \alpha T(t) + \beta \max\{0, t^* - t - T(t)\} + \gamma \max\{0, t + T(t) - t^*\} + cM(t), \tag{5}$$

where $\gamma > \alpha > \beta$ in accordance with empirical results (Small, 1982). The first term describes the travel time cost, the second time describes the early arrival cost, the third time describes the late arrival cost, and the fourth term describes the mobility-based road user charge conform Eqn. (4),

$$cM(t) = \mu T(t) + \delta L. \tag{6}$$

The experienced travel time when departing at time instant t, $T(t)$, can be written as a sum of a fixed free-flow travel time and a time-dependent queuing delay. The free-flow travel time depends on link length L and free-flow speed V, while the queuing delay depends on the number of vehicles encountered in the queue when departing at time instant t, denoted by $Q(t)$, and bottleneck exit capacity s. Hence,

$$T(t) = \frac{L}{V} + \frac{Q(t)}{s}. \tag{7}$$

Substituting Eqn. (6) and Eqn. (7) into Eqn. (5) yields

$$C(t) = (\alpha + \mu)\frac{Q(t)}{s} + \beta \max\left\{0, t^* - t - \frac{L}{V} - \frac{Q(t)}{s}\right\} + \gamma \max\left\{0, t + \frac{L}{V} + \frac{Q(t)}{s} - t^*\right\} + C^0, \tag{8}$$

where C^0 is a minimum generalised travel cost defined as

$$C^0 = (\alpha + \mu)\frac{L}{V} + \delta L. \tag{9}$$

We wish to determine equilibrium departure rates $r(t)$, where $\int_t r(t)dt = N$, such that no traveller can decrease their generalised travel cost by unilaterally changing departure time. It can be shown that these equilibrium departure rates are (see Appendix A for more details):

$$r(t) = \begin{cases} \left(\dfrac{\alpha+\mu}{\alpha+\mu-\beta}\right)s, & t \in [t', \tilde{t}), \\ \left(\dfrac{\alpha+\mu}{\alpha+\mu+\gamma}\right)s, & t \in [\tilde{t}, t''], \end{cases} \tag{10}$$

where

$$t' = t^* - \frac{L}{V} - \left(\frac{\gamma}{\beta+\gamma}\right)\left(\frac{N}{s}\right), \tag{11}$$

$$\tilde{t} = t^* - \frac{L}{V} - \left(\frac{\beta}{\alpha+\mu}\right)\left(\frac{\gamma}{\beta+\gamma}\right)\left(\frac{N}{s}\right), \tag{12}$$

$$t'' = t^* - \frac{L}{V} + \left(\frac{\beta}{\beta+\gamma}\right)\left(\frac{N}{s}\right). \tag{13}$$

As is well-known in the bottleneck model, these departure rates lead to a queue growing and dissipating over time due to the fact that most travellers prefer to arrive close to their preferred arrival time. The first and last travellers experience free-flow travel time T^0 but face large schedule delays, whereas the traveller that arrives exactly on-time experiences the largest queue and travel time (see e.g., Arnott et al., 1993). Cumulative departures and arrivals in the bottleneck model are visualised in Figure 8. In Figure 8 also the area that represents the total queuing delay is shaded, which represents the total congestion delay experienced by all travellers. This total congestion delay, denoted by H, can be computed as (see Appendix A for details):

$$H = \frac{1}{2}\left(\frac{\beta}{\alpha+\mu}\right)\left(\frac{\gamma}{\beta+\gamma}\right)\left(\frac{N^2}{s}\right). \tag{14}$$

Several observations can be made. First, as can be observed in Eqn. (10), travel time charging rate μ adds to the value of time α, and hence mobility-based charging makes travellers more sensitive to travel time such that more travellers start travelling earlier and later, resulting in a decrease in total congestion delay H. As shown in Figure 9, the total queuing delay becomes smaller with increasing travel time charging rate μ, where $H \to 0$ if $\mu \to \infty$. This means that mobility-based charging will always result in less delays than distance-based charging, although the size of the benefit will depend on the parameter values. Secondly, since δ has no influence on departure rates $r(t)$ or total queuing delays H, distance-based charging cannot mitigate congestion in this case. Thirdly, it is easy to see that the total mobility consumption (i.e., road space use) is equal to $\lambda(NL/V + H) + \tau NL$ and hence is directly proportional to H since all other components are fixed.

5.2 Impact on route choice

The objective of this example is to understand how mobility consumption charging affects route choice, i.e., which routes drivers are taking and how the equilibrium state is affected. Consider the network with two routes shown in Figure 10.

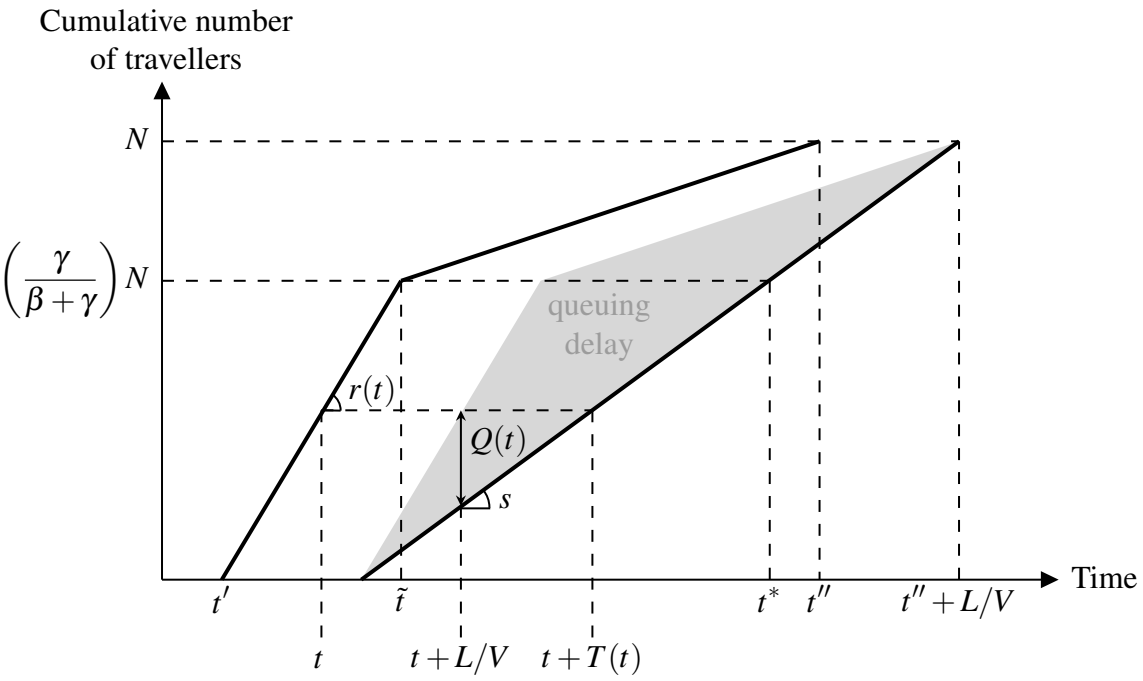

Figure 8: Cumulative departures and arrivals in bottleneck model

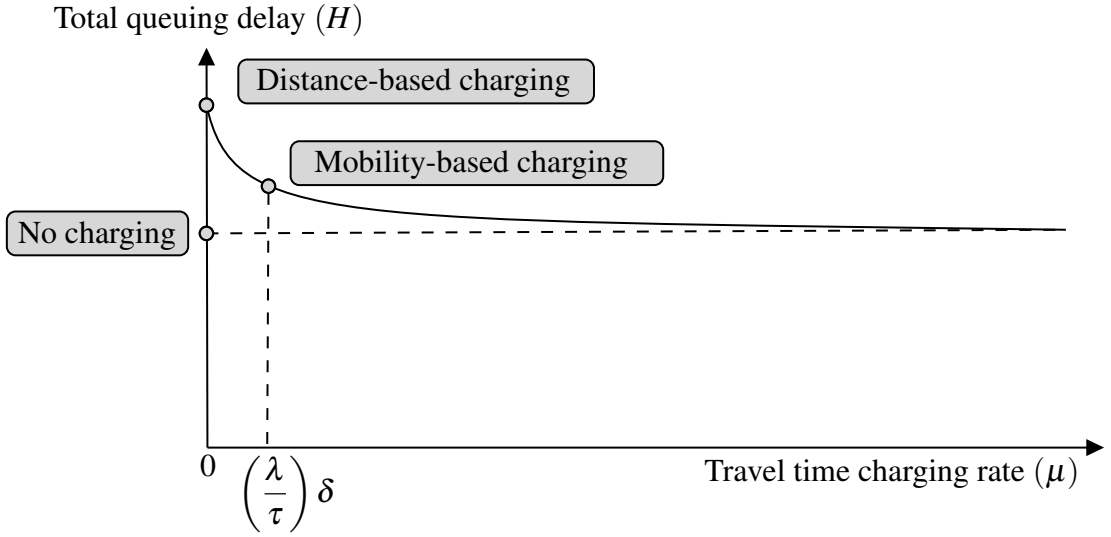

Figure 9: Total queuing delay depending on the travel time charging rate in the bottleneck model. This is an illustrative graph based on $N = 2000$, $s = 1000$, $\alpha = 18$, $\beta = 9$, and $\gamma = 36$. The shape of the graph is similar for other values.

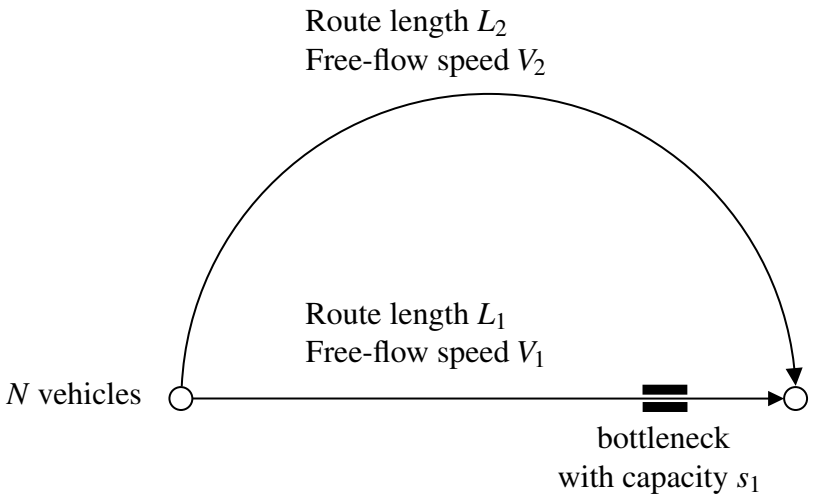

Figure 10: Dual route network.

Consider N travellers departing during time period $[t', t'']$ at a constant rate $r^* = N/(t'' - t')$ choosing their optimal route. Route 1 has length L_1 and free-flow speed V_1, while route 2 is longer with $L_2 > L_1$ but has a higher speed $V_2 > V_1$. Further we assume that route 1 has limited capacity $s_1 < r^*$, whereas route 2 has a high capacity $s_2 > r^*$. This is a common situation, for example route 1 can be considered an urban route through the city, while route 2 can be considered beltway around the city. The generalised travel cost for each route is given as

$$C_i(t) = \alpha T_i(t) + cM_i(t), \tag{15}$$

where the first term describes the travel time cost and the second term describes the mobility-based road user charge. Similar to the assumptions in Section 5.1, the mobility-based charge is conform Eqn. (4),

$$cM_i(t) = \mu T_i(t) + \delta L_i, \tag{16}$$

where the travel time on route i when departing at time t is the sum of a fixed free-flow travel time and a possible queuing delay that depends on the number of vehicles encountered in the queue, $Q_i(t)$, and bottleneck exit capacity s_i. The experienced travel time $T_i(t)$ when departing at time t on route i is given by

$$T_1(t) = \frac{L_1}{V_1} + \frac{Q_1(t)}{s_1}, \qquad \text{and} \qquad T_2(t) = \frac{L_2}{V_2}, \tag{17}$$

noting that route 2 will be uncongested since we assumed $s_2 > r^*$. Substituting Eqns. (16) and (17) into Eqn. (15) yields

$$C_1(t) = (\alpha + \mu)\frac{Q_1(t)}{s_1} + C_1^0, \qquad \text{and} \qquad C_2(t) = C_2^0. \tag{18}$$

where C_i^0 is the minimum generalised travel cost on route i, defined by

$$C_i^0 = (\alpha+\mu)\frac{L_i}{V_i} + \delta L_i. \tag{19}$$

Let the departure rate on route i at time t be denoted by $r_i(t)$. Here we only consider route choice and not departure time choice, so for flow conservation it needs to hold that

$$r_1(t) + r_2(t) = r^*, \qquad \text{for all} \quad t \in [t', t'']. \tag{20}$$

As a dynamic extension of Wardrop's equilibrium law (Wardrop, 1952), for each departure time $t \in [t', t'']$ we wish do determine equilibrium departure rates $r_i(t)$ such that no traveller can decrease their generalised travel cost by unilaterally changing route. Since $r_2(t)$ can be derived from $r_1(t)$ via Eqn. (20), we focus on determining equilibrium departure rates for route 1.

Let $\tilde{T}$ be the initial travel cost advantage of route 1, expressed in hours,

$$\tilde{T} = \frac{C_2^0 - C_1^0}{\alpha+\mu} = \frac{L_2}{V_2} - \frac{L_1}{V_1} + \left(\frac{\delta}{\alpha+\mu}\right)(L_2 - L_1). \tag{21}$$

The larger the difference in free-flow travel time and the larger the difference in travel distance, the larger $\tilde{T}$ will be, i.e., the longer it takes for the travel cost of route 1 to reach C_2^0. Three cases can be distinguished:

I) $\tilde{T} \leq 0$: Only route 2 is used, $r_1(t) = 0$ for all $t \in [t', t'']$;

II) $0 < \tilde{T} < \left(\frac{r^* - s_1}{s_1}\right)\left(\frac{N}{r^*}\right)$: Both routes are used, $r_1(t) \in (0, r^*)$ for all $t \in [t', t'']$;

III) $\tilde{T} \geq \left(\frac{r^* - s_1}{s_1}\right)\left(\frac{N}{r^*}\right)$: Only route 1 is used, $r_1(t) = r^*$ for all $t \in [t', t'']$.

Let $\tilde{t} \in [t', t'']$ denote the first moment that route 2 is used, which corresponds to the time instant when the queuing delay on route 1 reaches $\tilde{T}$. It can be shown that the equilibrium departure rates for route 1 are (see Appendix B),

$$r_1(t) = \begin{cases} r^*, & t \in [t', \tilde{t}\,], \\ \\ s_1, & t \in [\tilde{t}, t'']\,, \end{cases} \tag{22}$$

where

$$\tilde{t} = \begin{cases} t', & \text{in case I,} \\ t' + \left(\frac{s_1}{r^* - s_1}\right)\tilde{T}, & \text{in case II,} \\ t'', & \text{in case III.} \end{cases} \tag{23}$$

Cumulative departures and arrivals in the route choice model for the case where both routes 1 and 2 are used are visualised in Figure 11. In this figure, the total queuing delay, is indicated by

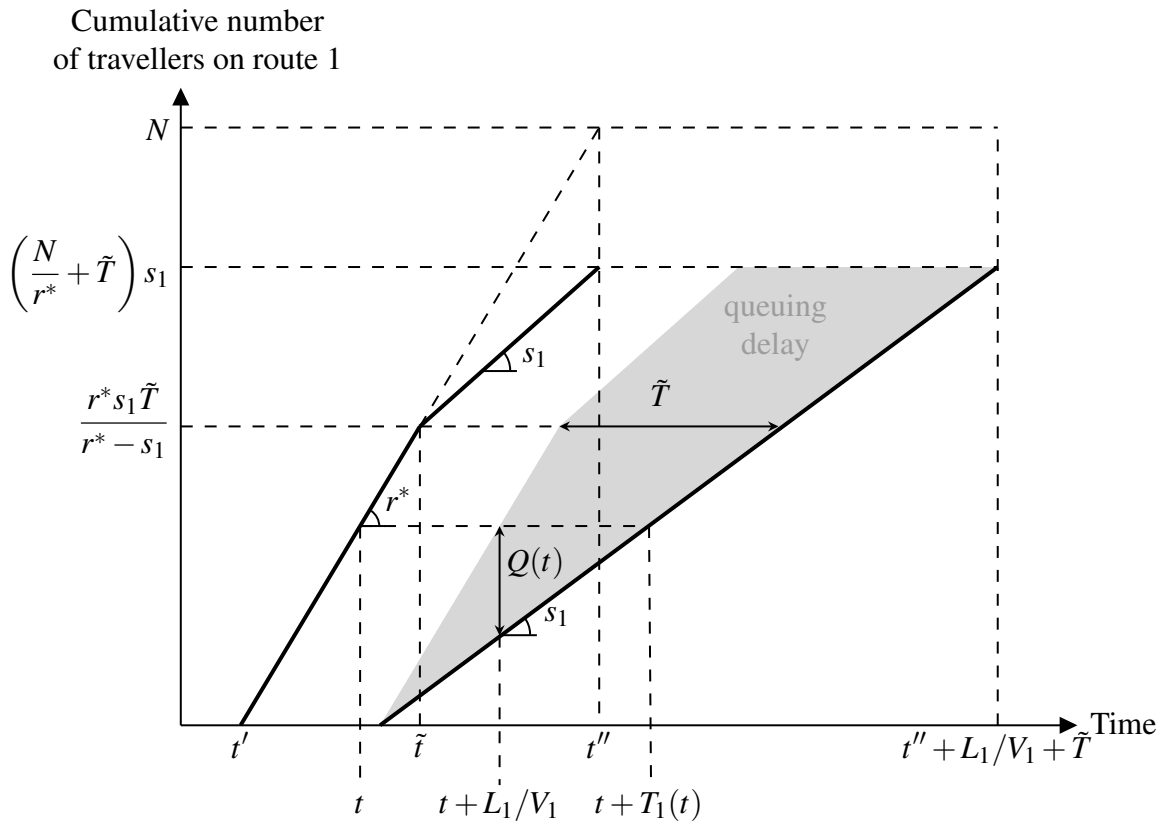

Figure 11: Cumulative departures and arrivals in route choice model.

shading in grey. This total queuing delay, denoted by H, can be computed as (see Appendix B for details):

$$H=\begin{cases}0, & \text{in case I,}\\ \left(\dfrac{N}{r^*}+\dfrac{(r^*-2s_1)\tilde{T}}{2(r^*-s_1)}\right)\tilde{T}s_1, & \text{in case II,}\\ \left(\dfrac{r^*-s}{2r^*s}\right)N^2, & \text{in case III.}\end{cases} \tag{24}$$

Several observations can be made. In the absence of road user charging, case I only occurs when the free-flow travel time on route 1 is larger than the free-flow travel time on route 2, and case III only occurs when the free-flow travel time on route 1 is much smaller than the free-flow travel time on route 2. Since $L_2 > L_1$, with distance-based charging ($\delta > 0$), route 1 becomes more attractive and therefore encourages more travellers to drive through the bottleneck. This is illustrated in Figure 11 where distance-based charging leads to more congestion compared to no charging. Mobility-based charging mitigates this effect by not only charging for distance but also encouraging travellers to take faster routes. While the total queuing delay decreases with increasing travel time charging rate, H only approaches the queuing delay that results from no charging when $\mu \to \infty$ and does not become zero.

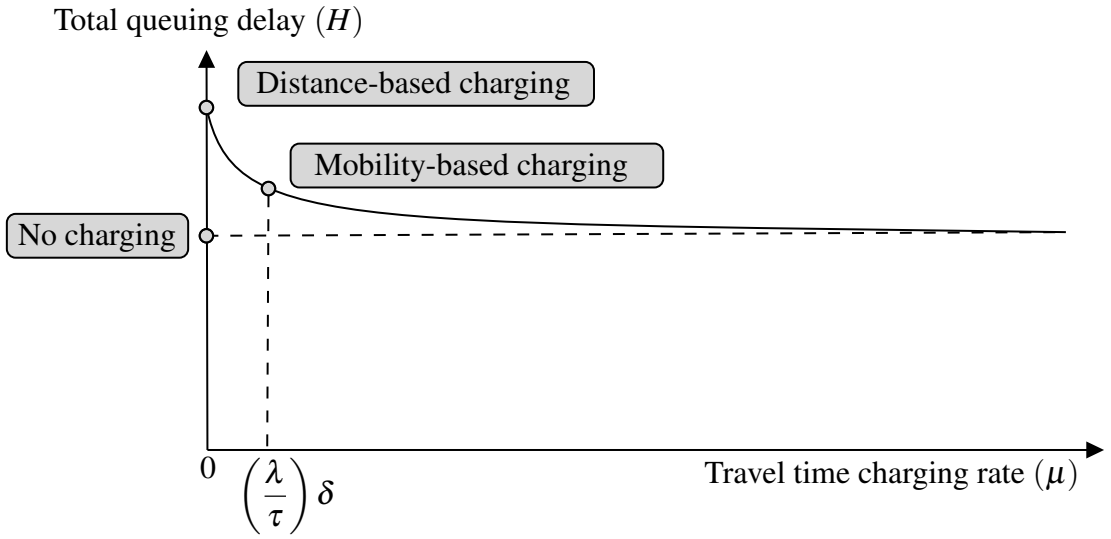

Figure 12: Total queuing delay depending on the travel time charging rate in case II. This is an illustrative graph based on $L_1 = 1$, $L_2 = 3$, $V_1 = 50$, $V_2 = 120$, $N = 2000$, $r^* = 2000$, $s_1 = 1000$, and $\alpha = 18$. The shape of the graph is similar for other values.

6 Technology innovation's impact on mobility consumption

In this section we discuss how mobility consumption and mobility-based charging is impacted by vehicle automation and connectivity (see Section 6.1) and electrification (see Section 6.2).

6.1 Connected and automated vehicles

According to Society of Automotive Engineers (SAE), there are 6 levels of vehicle automation, starting from Level 0 (no automation) to Level 5 (full automation) with various degrees of partial automation in between. Vehicle automation may affect safe space headways, especially if vehicle automation is complemented with vehicle-to-vehicle (V2V) connectivity that permits vehicle platooning at high speeds with a small gap between vehicles. Our proposed mobility consumption theory and mobility-based charging scheme are well equipped to capture the impact of connected and automated vehicles (CAVs) by simply adjusting physical parameters λ_n and τ_n.

When all vehicles on roads are connected and fully automated, applying our mobility consumption theory is straightforward. In this scenario, smaller values for λ_n and τ_n can be assigned since these vehicles can react more quickly and follow each other more closely than human-driven vehicles. In the literature, a typical value for the bumper-to-bumper spacing for CAVs is 2 m (Sun et al., 2020b), such that $\lambda_n = (4.87+2)/1000 = 0.00687$ for an average sized vehicle. With respect to reaction time, a typical value considered for CAVs is 0.5 s (Sun et al., 2018; Khoury et al., 2019), such that $\tau_n = 0.5/3600 = 0.000139$. Considering again the NGSIM I-80 data set that we analysed in Sections 3.2 and 3.3. Based on an average travel distance of 300 m and an average travel time of 28 s in this data set, CAVs could cut mobility consumption in half since for an average vehicle

the reduction in mobility consumption is

$$\frac{(6.25-2)\cdot 28+(1.23-0.5)\cdot 300}{11.12\cdot 28+1.23\cdot 300}\approx 0.50. \tag{25}$$

The situation is more complex in mixed traffic where human-driven vehicles co-exist with automated (but not necessarily connected) vehicles. Evidence in the literature suggests that in mixed traffic automated vehicles are likely to improve the safety performance at the cost of traffic efficiency by deliberately following other vehicles at a larger distance and cautiously responding to changes in the environment. That means, mobility consumption savings from automated vehicles in mixed traffic can be much smaller than expected, and can even be negative. For instance, Hu et al. (2023) extracted automated vehicles' response time using the Waymo Open Dataset, and found that at the current stage automated vehicles' response time is larger than human-driven vehicles' when operating on open public roads and displays significant variations as well, which is opposite to what has been commonly assumed in the literature. More specifically, the average response time for an automated vehicle following a human-driven vehicle across different traffic states is 2.5s, for a human-driven vehicle following an automated vehicle is 1.9s, and for a human-driven vehicle following a human-driven vehicle is 1.8 s. Thus, if the vehicles on I-80 were mixed with Waymo automated vehicles, the mobility consumption would increase by about 30%, assuming that the minimum space headway remains the same and that reaction time increases to 1.9 s. More information about the Waymo data can be found in Hu et al. (2022).

At different stages of vehicle automation and degrees of mixed traffic, adjusting the mobility-based charging rate c_n can effectively capture vehicle automation's complicated impact on traffic, which can be useful for assuaging people's resistance to allowing automated vehicles to operate in their city (Ariana Bindman, 2023).

6.2 Engine technologies

Mobility-based charging scheme can effectively mitigate the threat of government revenue loss caused by diminishing fuel excise tax revenue, arising from the electrification of vehicles by setting c_n to the values as illustrated in Section 4.2.

Moreover, as previously mentioned, the mobility consumption theory can also be extended to internalise external costs of traffic, e.g., CO_2 emissions. Adding the external cost dimension on top of the space and time dimension of the mobility consumption theory can be done at different levels, assuming that most external costs are related to travel time T_n or travel distance D_n. At the macroscopic level, for example, tailpipe CO_2 emissions per time unit, $e_n(t)$, for a vehicle n with an internal combustion engine are an almost linearly increasing function of speed $v_n(t)$ as shown in Barmpounakis et al. (2021). Thus, we can write

$$e_n = \eta_n^0 + \eta_n^1 v_n(t), \tag{26}$$

where η_n^0 and η_n^1 are vehicle-specific parameters. For electric vehicles it holds that $\eta_n^0 = \eta_n^1 = 0$. Similar to Eqn. (2), the total CO_2 emissions E_n can be calculated as

$$E_n = \int_{t=t_n^0}^{t_n^0+T_n} e_n(t)dt = \eta_n^0 T_n + \eta_n^1 D_n. \tag{27}$$

While the engine type does not affect mobility consumption, a government may want to provide incentives to reduce CO_2 emissions via a "polluter pays" carbon charging mechanism. Suppose that the carbon charging rate is w_n, then the combined mobility and carbon charge is equal to

$$c_n M_n + w_n E_n = \left(c_n \lambda_n + w_n \eta_n^0\right) T_n + \left(c_n \tau_n + w_n \eta_n^1\right) D_n = \mu_n^* T_n + \delta_n^* D_n. \tag{28}$$

In other words, it is straightforward to extend mobility-based charging with carbon charging by altering the travel time and travel distance charging rates. Monetary values for w_n can be found for example in European Commission (2019). Carbon charging can be linked with eco-driving systems (Sun et al., 2020a) to guide drivers to save energy and reduce emissions.

7 Discussion and conclusions

In this study, we made an analogy between mobility and electric energy and introduced a novel mobility consumption theory to measure road use based on "reserved" safe space headway over time. As a result, road use is measured in kilometre-hours, similar to power consumption measured in kilowatt-hours. Computing mobility consumption only requires measurements of travel time and travel distance (of a trip, month, or year), as well as vehicle-specific parameters.

We also illustrated how mobility consumption theory can be used for road user charging, which we distinctively defined as mobility-based charging. Mobility-based charging puts a price on each kilometre-hour of road use, similar to the way we pay for electricity. Since mobility-based charging boils down to a charge for travel distance and a charge for travel time, distance-based charging is a special case of mobility-based charging where travel time is not charged, while a specific form of congestion pricing results if one only charges for travel time and not for travel distance. A mobility-based charging scheme does not require an on-board GPS unit, instead only needs information from an existing on-board odometer and an additional meter that tracks cumulative running time of a vehicle, making it a non-intrusive way of road user charging. Since many people are already familiar with paying for time and distance in a taxi or via car sharing, mobility-based charging is transparent and easy to explain.

In short, mobility-based charging is a relatively simple road user charging scheme that is fair because it directly relates to road use, it provides a sustainable revenue sources because it charges for travel distance and travel time, and it is economically efficient because it not only charges more for driving more, but it also automatically charges more when driving in congested conditions. It can be extended beyond charges for road use to internalise external costs such as emissions. The proposed charging scheme could also be adopted within tradable mobility credits (Verhoef et al., 1997; Yang and Wang, 2011). Finally, mobility-based charging can naturally be extended to include further aspects of the transportation system such as parking fees that depend on the time a vehicle occupies road space.

As discussed in Section 6, mobility-based charging can be readily extended to effectively capture vehicle automation's complicated impact on traffic and include carbon charging. Other external costs could potentially also be internalised by charging for travel time and travel distance. Since crash risk increases with vehicle speed (International Transport Forum, 2018), a pay-as-you-

drive scheme that charges for both travel distance and travel time could be considered to reduce road trauma. Schröder et al. (2022) provides an overview of various other external mobility costs.

While we focused on the application of mobility consumption theory to road user charging, we believe that there is scope to also utilise it in the context of traffic and travel demand management. For example, it could be used to optimise road use at bottlenecks or at specific locations during events, or to increase efficiency of entire networks via variable speed limits, ramp metering, lane management, perimeter control, etc. There is also scope for mobility-based charging beyond car traffic as the principle can be applied to any mobile person or vehicle, e.g., buses, trains in rail networks, bikes on cycle lanes, etc.

We focused in this paper on core mechanisms of mobility consumption theory and mobility-based charging, but more research is required to further investigate real-world applicability and political feasibility. A desirable next step would be to conduct a field trial to test mobility-based charging in practice.

Acknowledgment

This project is partially funded by the Bavarian State Ministry of Science and the Arts in the framework of the bidt Graduate Center for Postdocs.

A Mathematical derivations for departure time choice model

Following Arnott et al. (1993), let $\tilde{t}$ be the departure time for on-time arrival, i.e., $\tilde{t}+T(\tilde{t})=t^*$, and let t' and t'' denote the departure times of the first and last traveller, respectively. We would like to determine rates $r(t)$ such that generalised travel costs are equal, i.e., $C(t)=\bar{C}$, for all $t\in[t',t'']$. The following equilibrium conditions should hold for early and late arrival, respectively:

$$\bar{C} = (\alpha+\mu)\frac{Q(t)}{s}+\beta\left(t^*-t-\frac{L}{V}-\frac{Q(t)}{s}\right), \qquad \text{for } t\in\left[t',\tilde{t}\right). \tag{29}$$

$$\bar{C} = (\alpha+\mu)\frac{Q(t)}{s}+\gamma\left(t+\frac{L}{V}+\frac{Q(t)}{s}-t^*\right), \qquad \text{for } t\in\left[\tilde{t},t''\right]. \tag{30}$$

Taking the first derivative yields

$$\frac{dQ(t)}{dt}=\frac{\beta s}{\alpha+\mu-\beta}, \quad t\in\left[t',\tilde{t}\right), \quad \text{and} \quad \frac{dQ(t)}{dt}=-\frac{\gamma s}{\alpha+\mu+\gamma}, \quad t\in\left[\tilde{t},t''\right]. \tag{31}$$

Since $Q(t)=\int_{\hat{t}}^{t} r(u)du-s(t-\hat{t})$, where $\hat{t}$ is the most recent departure time at which a traveller faced no queue, it holds that

$$\frac{dQ(t)}{dt}=r(t)-s. \tag{32}$$

Substitution of Eqn. (32) into Eqn. (31) and solving for $r(t)$ yields Eqn. (10). Values for t' and t'' can be determined by solving the following system of equations:

$$C(t'') - C(t') = \gamma\left(t'' + \frac{L}{V} - t^*\right) - \beta\left(t^* - t' - \frac{L}{V}\right) = 0, \qquad \text{and} \qquad t'' - t' = \frac{N}{s}. \tag{33}$$

The first equation needs to hold since $C(t'') = C(t') = \bar{C}$ and the second equation needs to hold because arrivals are continuous over time period $[t', t'']$ and exit flow rates are constant at capacity rate s. Solving for t' and t'' yields Eqns. (11) and (13). Using either of these time instants we can derive the equilibrium cost, which also needs to be equal to $C(\tilde{t})$, which results in

$$\bar{C} = \frac{\beta\gamma N}{(\beta+\gamma)s} = (\alpha+\mu)\frac{Q(\tilde{t})}{s}. \tag{34}$$

Further, the queue faced when departing at $\tilde{t}$ is equal to

$$Q(\tilde{t}) = \left(\frac{(\alpha+\mu)s}{\alpha+\mu-\beta} - s\right)(\tilde{t} - t') = \frac{\beta s}{\alpha+\mu-\beta}(\tilde{t} - t'). \tag{35}$$

Combining Eqns. (11), (35) and (34) and solving for $\tilde{t}$ yields Eqn. (12).

The total travel delay can be computed as the shaded area indicated in Figure 8,

$$H = \frac{NQ(\tilde{t})}{s}. \tag{36}$$

Substituting Eqn. (35) and

$$\tilde{t} - t' = \left(\frac{\gamma}{\alpha+\mu}\right)\left(\frac{\alpha+\mu-\beta}{\beta+\gamma}\right)\left(\frac{N}{s}\right), \tag{37}$$

yields Eqn. (14).

B Mathematical derivations for route choice model

Consider the initial travel cost advantage of route 1, expressed in hours, $\tilde{T}$ in Eqn. (21). In case I, $\tilde{T} \leq 0$, it holds that $C_2^0 \leq C_1^0$ so that travellers use route 2 directly at the start at time t'. Since route 2 is uncongested, there is no incentive to switch to route 1 and hence in equilibrium $r_1(t) = 0$ for all $t \in [t', t'']$.

In cases II, $\tilde{T} > 0$, which means that initially travellers will use route 1 but will start using route 2 at time $\tilde{t}$ once the queuing delay on route 1 reaches a level such that $C_1(\tilde{t}) = C_2^0$. In equilibrium, that means that $r_1(t) = r^*$ for $t \in [t', \tilde{t})$ while $r_1(t) = s_1$ for $t \in [\tilde{t}, t'']$ in order to stablise the queuing delay, which results in Eqn. (22). After time instant $\tilde{t}$ both routes are used and the associated equilibrium travel cost $\bar{C}$ is

$$\bar{C} = C_1(t) = (\alpha+\mu)\frac{Q_1(t)}{s_1} + C_1^0 = C_2^0, \qquad \text{for } t \in [\tilde{t}, t''']. \tag{38}$$

In other words, in equilibrium, the queuing delay experienced when departing at time $t \in [\tilde{t}, t'']$ is

$$\frac{Q_1(t)}{s_1} = \frac{C_2^0 - C_1^0}{\alpha + \mu} = \tilde{T}. \tag{39}$$

The number of vehicles in the queue on route 1 when departing at time instant t is equal to $Q_1(t) = \int_{t'}^{t} r(u)du - s_1(t' - t)$. Assuming that $\tilde{t} \in [t', t'']$, the number of vehicles in the queue when departing at time $\tilde{t}$ will according to Eqn. (39) reach an equilibrium when

$$Q_1(\tilde{t}) = (r^* - s_1)\left(\tilde{t} - t'\right) = \tilde{T} s_1. \tag{40}$$

This results in the second part of Eqn. (23),

$$\tilde{t} = t' + \left(\frac{s_1}{r^* - s_1}\right)\tilde{T}. \tag{41}$$

In case III, the queue on route 1 never gets long enough for travellers to divert to route 2, which means that $\tilde{t} \notin [t', t'']$. Therefore, Eqn. (41) only holds if $\tilde{t} < t''$, which means that it should hold that

$$t'' - \tilde{t} = t'' - t' - \left(\frac{s_1}{r^* - s_1}\right)\tilde{T} > 0. \tag{42}$$

Rewriting Eqn. (42) and remembering that $r^* = N/(t'' - t')$, we obtain

$$\tilde{T} < \left(\frac{r^* - s_1}{s_1}\right)\left(\frac{N}{r^*}\right). \tag{43}$$

If this condition does not hold, then $r_1(t) = r^*$ for the entire period $[t, t'']$, i.e., we simply set $\tilde{t} = t''$ in case III.

Consider case II. At time instant $\tilde{t}$, the cumulative number of travellers on route 1 equals $(\tilde{t} - t')r^*$. Using Eqn. (41), this yields $r^* s_1 \tilde{T}/(r^* - s_1)$. The first of these travellers experiences a delay of zero, while the last experiences a queuing delay of $\tilde{T}$, hence in total these travellers experience a delay of $\frac{1}{2} r^* s_1 \tilde{T}^2/(r^* - s_1)$. During time period $[\tilde{t}, t'']$ an additional $(t'' - \tilde{t})s_1$ travellers take route 1, which is equal to $N s_1/r^* - s_1^2 \tilde{T}/(r^* - s_1)$. Each of these travellers experiences a queuing delay of $\tilde{T}$. Therefore, the total queuing delay is

$$H = \frac{r^* s_1 \tilde{T}^2}{2(r^* - s_1)} + \left(\frac{N}{r^*} - \frac{s_1 \tilde{T}}{r^* - s_1}\right)\tilde{T} s_1 = \left(\frac{N}{r^*} + \frac{(r^* - 2s_1)\tilde{T}}{2(r^* - s_1)}\right)\tilde{T} s_1, \tag{44}$$

which establishes Eqn. (24) for case II. For case I, it is obvious that $H = 0$. To compute the total queuing delay for case III we substitute the upper bound on $\tilde{T}$ from Eqn. (43) in Eqn. (44), which yields Eqn. (24) for case III.

References

Ariana Bindman, S., 2023. 'Incompetent' driverless cars are wreaking havoc on San Francisco — sfgate.com. https://www.sfgate.com/tech/article/cruise-waymo-driverless-cars-san-francisco-18132953.php. [Accessed 17-Jun-2023].

Arnott, R., de Palma, A., Lindsey, R., 1993. A Structural Model of Peak-Period Congestion: A Traffic Bottleneck with Elastic Demand. The American Economic Review 83, 161–179. URL: http://www.jstor.org/stable/2117502.

Australian Automobile Association, 2019. Road congestion in Australia. Technical Report. https://www.aaa.asn.au/wp-content/uploads/2019/06/Road-Congestion-In-Australia-2019-v.3.pdf.

Australian Bureau of Statistics, 2019. Survey of motor vehicle use. Technical Report. Canberra ACT, Australia.

Barmpounakis, E., Montesinos-Ferrer, M., Gonzales, E.J., Geroliminis, N., 2021. Empirical investigation of the emission-macroscopic fundamental diagram. Transportation Research Part D: Transport and Environment 101, 103090. doi:10.1016/J.TRD.2021.103090. publisher: Pergamon.

Daganzo, C.F., 2007. Urban gridlock: Macroscopic modeling and mitigation approaches. Transportation Research Part B: Methodological 41, 49–62. doi:10.1016/j.trb.2006.03.001.

European Commission, 2019. Handbook on the external costs of transport. Publications Office, LU. URL: https://data.europa.eu/doi/10.2832/51388.

Green, M., 2000. "How long does it take to stop?" Methodological analysis of driver perception-brake times. Transportation Human Factors 2, 195–216. doi:10.1207/STHF0203_1.

Greenshields, B.D., 1935. A study of highway capacity. Highway Research Board Proceedings 14, 448–477.

Helbing, D., 2001. Traffic and related self-driven many-particle systems. Reviews of Modern Physics 73, 1067–1141. doi:10.1103/RevModPhys.73.1067.

Hensher, D.A., 2020. Electric cars – they may in time increase car use without effective road pricing reform and risk lifecycle carbon emission increases. Transport Reviews 40, 265–266. doi:10.1080/01441647.2020.1709273.

Hensher, D.A., Bliemer, M.C., 2014. What type of road pricing scheme might appeal to politicians? Viewpoints on the challenge in gaining the citizen and public servant vote by staging reform. Transportation Research Part A: Policy and Practice 61, 227–237. doi:10.1016/j.tra.2014.02.017.

Hensher, D.A., Mulley, C., 2014. Complementing distance based charges with discounted registration fees in the reform of road user charges: the impact for motorists and government revenue. Transportation 41, 697–715. doi:10.1007/s11116-013-9473-6.

Herman, R., Prigogine, I., 1979. A two-fluid approach to town traffic. Science 204, 148–151.

Hu, X., Zheng, Z., Chen, D., Sun, J., 2023. Autonomous vehicle's impact on traffic: empirical evidence from waymo open dataset and implications from modelling. IEEE Transactions on Intelligent Transportation Systems 24, 6711–6724. doi:10.1109/TITS.2023.3258145.

Hu, X., Zheng, Z., Chen, D., Zhang, X., Sun, J., 2022. Processing, assessing, and enhancing the waymo autonomous vehicle open dataset for driving behavior research. Transportation Research Part C: Emerging Technologies 134, 103490. doi:https://doi.org/10.1016/j.trc.2021.103490.

International Transport Forum, 2018. Speed and Crash Risk. Technical Report. Paris.

Khoury, J., Amine, K., Abi Saad, R., 2019. An initial investigation of the effects of a fully automated vehicle fleet on geometric design. Journal of Advanced Transportation 2019, 6126408. doi:10.1155/2019/6126408.

Konstantinou, T., Labi, S., Gkritza, K., 2022. Assessing highway revenue impacts of electric vehicles using a case study. Research in Transportation Economics , 101248doi:10.1016/j.retrec.2022.101248.

Li, X., Peng, F., Ouyang, Y., 2010. Measurement and estimation of traffic oscillation properties. Transportation Research Part B: Methodological 44, 1–14. doi:10.1016/j.trb.2009.05.003.

Lindsey, R., Santos, G., 2020. Addressing transportation and environmental externalities with economics: Are policy makers listening? Research in Transportation Economics 82, 100872. doi:10.1016/j.retrec.2020.100872.

Ma, T., Ahn, S., 2008. Comparisons of Speed-Spacing Relations under General Car following versus Lane Changing. Transportation Research Record: Journal of the Transportation Research Board 2088, 138–147. doi:10.3141/2088-15.

Metz, D., 2018. Tackling urban traffic congestion: The experience of London, Stockholm and Singapore. Case Studies on Transport Policy 6, 494–498. doi:10.1016/j.cstp.2018.06.002.

Newell, G.F., 1993. A simplified theory of kinematic waves in highway traffic, part I: General theory. Transportation Research Part B: Methodological 27, 281–287.

Newell, G.F., 2002. A simplified car-following theory: a lower order model. Transportation Research Part B: Methodological 36, 195–205. doi:10.1016/S0191-2615(00)00044-8.

de Palma, A., Lindsey, R., 2011. Traffic congestion pricing methodologies and technologies. Transportation Research Part C: Emerging Technologies 19, 1377–1399. doi:10.1016/j.trc.2011.02.010.

Pigou, A.C., 1920. The economics of welfare. Macmillan, London.

Pipes, L.A., 1953. An Operational Analysis of Traffic Dynamics. Journal of Applied Physics 24, 274–281. doi:10.1063/1.1721265.

Queensland Government, 2022. Rules for heavy vehicles. https://www.qld.gov.au/transport/safety/heavy/rules.

Ranjitkar, P., Nakatsuji, T., Azuta, Y., Gurusinghe, G.S., 2003. Stability analysis based on instantaneous driving behavior using car-following data. Transportation Research Record 1852, 140–151. doi:10.3141/1852-18.

Rigot-Müller, P., 2018. Analysing the heavy goods vehicle "écotaxe" in France: Why did a promising idea fail in implementation? Transportation Research Part A: Policy and Practice 118, 147–173. doi:10.1016/j.tra.2018.08.024.

Schröder, D., Kirn, L., Kinigadner, J., Loder, A., Lienkamp, M., 2022. Ending the myth of mobility at zero costs: An external cost analysis. Research in Transportation Economics 97, 101246. doi:10.1016/j.retrec.2022.101246.

Sharma, A., Zheng, Z., Kim, J., Bhaskar, A., Haque, M.M., 2109. Estimating and Comparing Response Times in Traditional and Connected Environments. Transportation Research Record: Journal of the Transportation Research Board 2673, 674–684. doi:10.1177/0361198119837964.

Small, K.A., 1982. The Scheduling of Consumer Activities: Work Trips. The American Economic Review 73, 467–479.

Small, K.A., Verhoef, E.T., 2007. The economics of urban transportation. Routledge, London.

Steinsland, C., Fridstrøm, L., Madslien, A., Minken, H., 2018. The climate, economic and equity effects of fuel tax, road toll and commuter tax credit. Transport Policy 72, 225–241. doi:10.1016/j.tranpol.2018.04.019.

Sun, C., Guanetti, J., Borrelli, F., Moura, S.J., 2020a. Optimal eco-driving control of connected and autonomous vehicles through signalized intersections. IEEE Internet of Things Journal 7, 3759–3773. doi:10.1109/JIOT.2020.2968120.

Sun, J., Zheng, Z., Sun, J., 2018. Stability analysis methods and their applicability to car-following models in conventional and connected environments. Transportation Research Part B: Methodological 109, 212–237. doi:10.1016/j.trb.2018.01.013.

Sun, J., Zheng, Z., Sun, J., 2020b. The relationship between car following string instability and traffic oscillations in finite-sized platoons and its use in easing congestion via connected and automated vehicles with IDM based controller. Transportation Research Part B: Methodological 142, 58–83. doi:10.1016/j.trb.2020.10.004.

U.S. Department of Transportation Federal Highway Administration, 2016. Next Generation Simulation (NGSIM) Vehicle Trajectories and Supporting Data. Provided by ITS DataHub through Data.transportation.gov. doi:10.21949/1504477.

Verhoef, E., Nijkamp, P., Rietveld, P., 1997. Tradeable Permits: Their Potential in the Regulation of Road Transport Externalities. Environment and Planning B: Planning and Design 24, 527–548. doi:10.1068/b240527.

Verhoef, E.T., Bliemer, M.C.J., Steg, L., van Wee, B. (Eds.), 2008. Pricing in road transport: a multi-disciplinary perspective. Edward Elgar, Cheltenham, UK; Northampton, MA.

Vickrey, W.S., 1969. Congestion theory and transportation investment. The American Economic Review 59, 251–260. doi:10.1126/science.151.3712.867-a.

Vickrey, W.S., 1993. Point of view: principles and applications of congestion pricing. TR News 167, 4–5.

van Wageningen-Kessels, F., van Lint, H., Vuik, K., Hoogendoorn, S., 2015. Genealogy of traffic flow models. EURO Journal on Transportation and Logistics 4, 445–473. doi:10.1007/s13676-014-0045-5.

Wardrop, J.G., 1952. Some Theoretical Aspects of Road Traffic Research. Proceedings of the Institution of Civil Engineers Part II, 325–378.

Yang, H., Wang, X., 2011. Managing network mobility with tradable credits. Transportation Research Part B: Methodological 45, 580–594. doi:10.1016/j.trb.2010.10.002.

Zheng, Z., Ahn, S., Chen, D., Laval, J., 2011a. Applications of wavelet transform for analysis of freeway traffic: Bottlenecks, transient traffic, and traffic oscillations. Transportation Research Part B: Methodological 45, 372–384. doi:10.1016/j.trb.2010.08.002.

Zheng, Z., Ahn, S., Chen, D., Laval, J., 2011b. Freeway traffic oscillations: microscopic analysis of formations and propagations using wavelet transform. Procedia-Social and Behavioral Sciences 17, 702–716. doi:10.1016/j.sbspro.2011.04.540.

Zheng, Z., Ahn, S., Chen, D., Laval, J., 2013. The effects of lane-changing on the immediate follower: Anticipation, relaxation, and change in driver characteristics. Transportation Research Part C: Emerging Technologies 26, 367–379. doi:10.1016/j.trc.2012.10.007.