Congestion Pricing: A Mechanism Design Approach

C.-Philipp Heller [*] Johannes Johnen [†] Sebastian Schmitz [‡]

May 14, 2018

Abstract

We study traffic congestion as a mechanism design problem by analyzing the allocation of drivers to a congestible road. Drivers have private information about their value of time (VOT). With a finite number of drivers, the efficient allocation depends on drivers' VOTs and is ex-ante unknown. Thus, setting a single Pigouvian price is generally not optimal. Nevertheless, the regulator can implement the efficient allocation with a Vickrey-Clarke-Groves payment rule: drivers pay for road access but also for *faster travel*. Our mechanism sets this price correctly without prior knowledge of the distribution of drivers' VOT.

JEL classification: D82, D62, R48, R41
Keywords: Mechanism Design, Congestion Pricing, VCG Mechanism, Externalities, Value of Time

[*] NERA Economic Consulting, Germany, philipp.heller@nera.com

[†] Université catholique de Louvain and CORE, Voie Roman Pays 34, L1.03.01 B-1348 Louvain-la-Neuve, Belgium johannes.johnen@uclouvain.be

[‡] Federal Ministry for Economic Affairs and Energy, Germany, sebastian.schmitz@bmwi.bund.de

We thank Andreas Asseyer, Helmut Bester, Inacio Bo, Jan Ehmke, Veronika Grimm, Paul Heidhues, Johannes Hörner, Sebastian Kohls, Rafael Mourão, Shaofang Qi, Roland Strausz, Achim Wambach, Georg Weizsäcker and participants of the Microeconomic Theory Colloquium at HU and FU Berlin, the SFB/TR Workshop in Bonn, the SED Conference in Istanbul and EEA 2015 in Mannheim for valuable comments. We also thank two anonymous referees. Their comments improved the paper considerably. Financial support from the Fonds Spéciaux de Recherche (Johnen) and the DFG is gratefully acknowledged (Heller, Johnen). The usual disclaimer applies. Opinions expressed are solely our own and do not reflect the views or opinions of our employers.

1 Introduction

Congestion is an ever-present nuisance to many people living in the world's urban areas and has significant economic costs. Congestion adds about 5 cents per mile to the cost of travel, making it the single largest source of traffic-related externalities (Parry et al. 2007).[1] Schrank et al. (2015) estimate delay and fuel costs of $160bn due to congestion for the USA in 2014. From an economist's point of view, congestion is an externality problem since an individual driver does not take the effect of her journey on other drivers' travel time into account.

The classic solution is a Pigouvian tax set at a level to ensure that each driver internalizes the marginal cost of the increased travel time of other drivers. Much of the existing literature on congestion pricing is concerned with finding the right level of such a congestion charge given observable characteristics.[2] In practice this requires the regulator to have a reliable estimate of the VOT. Generally one would expect these values to differ substantially across drivers, even for the same driver, depending on trip purpose or the time of the day. Indeed Small et al. (2005) as well as Steimetz and Brownstone (2005) have empirically demonstrated the existence of substantial heterogeneity in the VOT from both observed and unobserved sources. This prevalence of unobserved heterogeneity suggests that traffic regulation could be improved if one could find a way to observe valuations of travel time directly, without trying to estimate them from observable characteristics.

In this paper we propose an application of mechanism design as a novel approach to deal with demand uncertainty in traffic regulation. Using mechanism design allows us to target three important issues. First, we can analyze the implications of the regulator's lack of detailed knowledge concerning the VOT. Second, we can use more general payment schemes that do not impose specific functional forms, such as a uniform Pigouvian price. Third, we obtain a better understanding about the conditions under which a Pigouvian price is appropriate to deal with traffic congestion.

[1] More precisely, 5 cents per mile is the marginal external cost of congestion of a representative sample of urban and rural roads in the US at different times of the day, weighted by vehicle miles traveled. Table 2 in Parry et al. (2007) summarizes the costs of traffic-related externalities. The estimated external cost of local pollution caused by automobiles is 2 cents per mile and 3 cents per mile for accident-related externalities. Externalities of greater greenhouse gas emissions is 0.3 cents per mile.

[2] Some examples of observable characteristics are time of day, income, measures of traffic density etc.

In our model each driver has private information about her VOT. To illustrate our results we study a travel-demand problem where drivers can chose to travel on a congestible road or not.[3] The externality of a driver on the others arises from slowing down other traveling drivers. A mechanism designer, who can be seen as a local government authority, a regulator, or a private company administering traffic, can design mechanisms involving pricing schemes and allocations to implement efficient road usage.[4]

We characterize the efficient allocation in this model. For a finite number of drivers, the efficient allocation depends on the *realized* VOTs. Since these realized values are ex-ante unknown to the traffic regulator, also the efficient allocation is ex-ante uncertain. Drivers reveal their travel preferences to a regulator who then allocates drivers efficiently to traveling or not traveling. We derive a payment scheme which ensures that drivers have the right incentives to reveal their travel preferences truthfully, and—with these information—implements the efficient allocation. This payment scheme is an application of the well-known Vickrey-Clarke-Groves (VCG) mechanism[5] and reflects the externality that the presence of the driver imposes on others.

Intuitively, our mechanism differs from a classic Pigouvian price in the following way: a Pigouvian price allows drivers to pay for road usage, i.e. traveling. If, however, many drivers have sufficiently large VOTs, they will pay this price but the resulting allocation makes the road inefficiently slow. Conversely, when VOTs are relatively low for most drivers, a single Pigouvian price that does not adjust immediately to these realizations leads to an underused road which is inefficiently fast. This illustrates that the efficient allocation depends on the realized VOTs. Our mechanism deals with these issues in two ways. First, in addition to paying for usage of the road, our mechanism allows drivers to reduce travel time on the road by making additional payments. Second, if

[3] One can think of drivers as commuters who decide whether to travel by car or take an alternative, such as public transport or working from home. We will refer to drivers as *traveling* or *not traveling* to indicate road usage. We focus on the travel-demand problem in the main text to simplify exposition. This travel-demand problem is equivalent to a road-choice problem where the outside option is an uncongestible road. A similar model is studied in Mayet and Hansen (2000). We study a road-choice problem with two congestible roads in the supplementary online material.

[4] In the supplementary material, we also consider a mechanism designer who wants to maximize revenue.

[5] This mechanism is inspired by the seminal works of Vickrey (1961), Clarke (1971) and Groves (1973). VCG-type mechanisms are used in practice for example by facebook to sell advertisement space in their AdAuction system, or in ebay auctions.

many drivers have relatively low VOTs, our mechanism learns about that by the drivers' reports and adjusts the price to enter the road, inducing efficient road usage in this case as well. In equilibrium, revealing your VOT is analogous to deciding out of a menu of desired travel time-price combinations. We will illustrate this in some examples in Section 3.

Our procedure opens the possibility of using second-degree price discrimination *within* a congestible road to extract drivers' VOT.[6] Even if the regulator does not have any prior knowledge about the distribution of the VOTs, our mechanism instantaneously provides the real time valuations of drivers. Thus, compared to using estimated VOTs our approach is flexible and adaptive. Regulators can deal with aggregate uncertainty and can respond quickly to changes in demand.

This result has important implications for the optimal pricing structure of congestion pricing. We show that even when the regulator knows the exact distribution of the VOTs, a single Pigouvian tax does not generally implement the efficient allocations when the number of drivers is finite. Previous models of congestion pricing took as granted a given pricing structure for the regulator. Each driver faces a price that could depend on observables, such as which road was used, time of the day, type of car, current or previous speed-flow on the road etc. Conditional on these observables the previous models could only look for the optimal price level within the pricing structure they imposed. Using mechanism design allows us to solve simultaneously for the optimal pricing structure, e.g. whether to use a single Pigouvian price or price discrimination, and the optimal prices within that pricing structure. This enables the regulator to also take unobservable characteristics related to the VOTs into account. Applying mechanism design allows us to show that a pricing structure that conditions on observables is not optimal in general. We show that with a finite number of drivers, more sophisticated payment schemes are required to implement the efficient allocation.

Our mechanism might have seemed impractical in the past as there was no way to easily communicate the VOT to a central decision-maker. There was also no way for the decision maker to force drivers to take particular routes. Modern communication technology, such as smart phones with navigation- or ride-hailing apps and GPS, and the advent of self-driving cars imply that these

[6] The literature already considers price discrimination between roads. For an example see Mayet and Hansen (2000). We show that this approach is generally not sufficient to implement efficient allocations in a single road.

practical problems may soon be overcome.[7]

1.1 Related Literature

Our approach is related to the literature on congestion pricing, and mechanism design.

Up to our knowledge, we are the first who study congestion externalities as a mechanism design problem. The existing literature on road congestion either assumes Pigouvian taxes or focuses on cases with infinitely many "small" drivers where Pigouvian taxes are efficient. We establish that Pigouvian prices are not efficient in the presence of aggregate uncertainty, i.e. the case with a finite number of drivers. We characterize efficient payment schemes for these cases.

1.1.1 Policy design for road congestion with heterogeneous preferences

The idea to use prices to implement efficient road usage dates back to Pigou (1920) and Knight (1924). Later this approach gained popularity among economists, for example with the seminal work of Vickrey and Sharp (Vickrey and Sharp 1968, Vickrey 1969). The central idea of this literature, the *Pigouvian approach*, is to implement efficient road usage by internalizing the social cost of congestion via a tax or price regulation: the surcharge for road usage is set to equate the marginal social cost at the efficient level.

Subsequent empirical and theoretical research has identified several problems with this approach, namely information requirements, and the users' heterogeneity in their VOT across users but also for a single user over the course of the day (for an overview, see Small 2012). Based on data on the use of pay lanes Small et al. (2005) estimate the distribution of the VOT of commuters in California choosing between a tolled express lane and a free alternative lane. They find a median VOT of around $23 with substantial heterogeneity unrelated to observable factors. Steimetz and Brownstone (2005) use commuters' highway choices to characterize the heterogeneity in the VOT

[7] With self-driving cars travelers could simply indicate their desired destination and their desire for reaching it on time. A self-driving car would electronically transmit this information to a central authority, which calculates an efficient travel schedule for the self-driving cars. Singapore for instance is already experimenting with self-driving cabs (Land Transport Authority 2016b).

by observable characteristics. They find that while the mean VOT is $30 per hour, these values range from $7 to $65 per hour between users.

Several theoretical papers analyze the congestion pricing problem with heterogeneity on the value for the trip. For example, Bernstein and El Sanhouri (1994) and Verhoef et al. (1996) analyze the problem of optimally setting congestion charges in a network with two roads, in which only one road can be tolled. In our paper, in contrast, the heterogeneity of drivers concerns the VOT.

Another strand of the literature has studied constrained optimal congestion pricing when there is heterogeneity in the VOT. We call these pricing schemes constrained optimal because, in contrast to our approach, they impose a specific functional form on pricing schemes. Closely related to our paper is Mayet and Hansen (2000), who study a road-choice model that is very similar to the travel-demand model that we study in the main section. Like in our model the heterogeneity of drivers concerns valuation of travel time, rather than the value of a trip. However they restrict the regulator to setting a single toll for using the congestible road. Small and Yan (2001) consider a model in which there are only two types of drivers, one with a high VOT and another with a low VOT. They highlight that due to the heterogeneity, there is some welfare gain from having roads with different travel times, as drivers with a high VOT will be willing to pay more to reach their destination faster. Arnott et al. (1994) analyze the choice of an optimal time-varying toll in a model with a heterogeneous VOT and random departure times. The number of drivers of each type in this model is known ex ante. Arnott and Kraus (1998) distinguish between anonymous and non-anonymous congestion charges and investigate under which conditions an anonymous congestion charge is optimal, when drivers can have varying VOTs. Unlike our model, the departure times may vary across drivers.

Another strand of the traffic-congestion literature, such as Vickrey (1967) or Arnott et al. (1990) considers a different source of congestion in a bottleneck problem. In these models congestion pricing aims at influencing drivers' departure time to regulate traffic. Our approach is flexible enough to be applied to this travel-scheduling problem. In fact, one can interpret our extension with two identical roads in Section 6.1 as a basic travel-scheduling model with two departure times.

However, we leave a detailed investigation of this problem for future research.

One feature of those papers is that they impose simple Pigouvian taxes. In contrast, we do not impose a functional form on prices, allowing us to study the form of generally optimal pricing schemes. We study explicitly the role of aggregate uncertainty over the number of drivers on the road and over the optimal level of road usage. When setting a single fixed congestion charge the number of drivers using a road is not determined. It depends on the number of drivers that have a sufficiently large VOT. In contrast, our mechanism uses the drivers' revealed VOTs to precisely determine the optimal number of drivers on the road. In general a single congestion charge (for each observable driver type) is not optimal. This point has not been recognized in the earlier literature.

To bridge the gap with the existing literature, we also derive conditions under which the welfare loss of using a single Pigouvian price becomes negligible. In this way, we contribute to understanding the conditions under which a Pigouvian congestion tax is appropriate, and when it could be improved upon. A single congestion charge can implement efficient traffic if the regulator already knows the efficient allocation of traffic. However, if she is uncertain about the efficient allocation, she could use a more complex price schedule, allowing her to learn about the efficient allocation while implementing it at the same time.

Instead of estimating demand, regulators could use trial-and-error methods proposed for instance by Vickrey (1993) or Downs (1993).[8] Our mechanism is designed to enable even an uninformed traffic regulator to use price discrimination to induce drivers to reveal their travel preferences truthfully. In this way, our mechanism is in principle very flexible and able to adjust quickly to changes in traffic demand.

1.1.2 Mechanism design with externalities

Methodically, we build on the literature on mechanism design. Mechanism design is a useful tool to study problems with demand uncertainty. Examples of these include the sale of items (Myerson

[8]Intuitively, these methods use observable, realized speed-flow relationships to adjust existing pricing schemes. This leads to a trial-and-error algorithm that converges to socially optimal congestion charge. For a survey on the existing literature see Yang et al. (2005), or de Palma and Lindsey (2011)

1981) where agents have private information on their valuations of the item; or the provision of public goods (Clarke 1971) where agents have private information about their own valuations of the public good. This includes cases where agents have private information about their costs for reducing emissions (Montero 2008).

Jehiel et al. (1996, 1999) study auctions with externalities where each buyer has a payoff from owning the item as well as from others owning the item. Agents cannot escape the externality, i.e. even agents who do not participate in the auction suffer (or benefit) from it. In the context of road congestion,however, the externality only affects drivers on the congestible road. Another difference of our paper is that we do not restrict the mechanism designer to selling a single good. In principle, each agent can be assigned to the congestible road.

VCG-type mechanims have also been used to study efficient solutions to environmental externalities. Montero (2008) looks at the problem of emissions abatement where polluters are privately informed about their cost of abatement. Traditionally proposed solutions, such as a tax on emissions or an emissions trading scheme are not efficient mechanisms in this context. Montero (2008) proposes instead a VCG-type mechanism to give polluters an incentive to reveal their cost of abatement truthfully and to implement the efficient level of abatement.

We contribute to the literature on mechanism design by showing that existing methods such as the VCG mechanism are useful to study the regulation of road congestion. To the best of our knowledge, road congestion has not been studied before in the literature on mechanism design.

The following Section 2 introduces our basic model and solves for the efficient allocation. Section 3 derives a general payment schedule that implements the efficient allocation and discusses simple examples. Section 4 discusses the relation of our proposed mechanism with a Pigouvian price by looking at the limit case when the number of drivers becomes large, and presents results of a simulation. Section 5 concludes. In the supplementary online material 6.1 we extend our results to a lane-pricing model where drivers can choose between two congestible roads. We study the case when the mechanism designer maximizes revenue rather than welfare in 6.2, and discuss currently used congestion pricing schemes in light of our results in 6.3.

2 Model

There are n drivers that simultaneously want to reach some common destination D, starting from a common starting point O, and a mechanism designer. The mechanism designer can represent a local government or a toll authority. Drivers can decide to do the trip T or to opt out O. If a driver travels, she uses a road that becomes congested as more drivers use it. If the driver opts out, she gets the utility of her outside option.[9] On the road, the travel time increases with the number of drivers.

We assume that drivers differ by the value they attach to the time spent traveling on this trip. This information is summarized for each driver i in the parameter $\theta_i \in \Theta_i \subseteq \mathbb{R}_+ \backslash \{0\}$. For some results we assume additionally that all θ_i are independently and identically distributed according to the well-behaved cumulative distribution function $F(\theta_i)$.[10] For most of our analysis the assumption on the distribution of the VOT $F(\cdot)$ is not necessary as the efficient mechanism induces the revelation of each drivers' VOT independent of distributional assumptions. However the distribution of drivers' VOT is needed when we consider limit cases and revenue maximization. We let $\theta \in \Theta \equiv \times_i \Theta_i$ be the n-dimensional vector of all drivers' valuation of time. We assume that for all i, j, $\theta_i \neq \theta_j$. Given well-behaved distribution functions, this case is expected to occur with certainty. We will denote by θ_{-i} the vector of all valuations except that of driver i.[11]

We normalize the utility of a driver's outside option to be zero, i.e. $u_i^O(\theta_i, k, p) = 0$.[12] The

[9] We can equivalently interpret this travel-demand problem as a road-choice problem. In this interpretation the outside option represents traveling on a slower but uncongested alternative road. Assuming an uncongested alternative simplifies the exposition of results but is not crucial, as we show in the supplementary online material 6.1.

[10] Well-behaved means that $F(\cdot)$ is continuously differentiable with a strictly positive derivative $f(\cdot)$ and such that $\frac{1-F(\theta_i)}{f(\theta_i)}$ is weakly decreasing in θ_i.

[11] With a slight abuse of notation we denote by $F(\theta)$ the joint distribution of the vector of valuations and $F(\theta_{-i})$ the joint distribution of all valuations except for that of driver i.

[12] We can maintain incentives when charging also for the outside option, i.e. by a tax for owning a car. In the extension in the supplementary online material 6.1 where the outside option represents another congestible road, driving that road can require positive payments.

utility of a driver who travels is given by:

$$u_i^T(\theta_i, k, p) = v(\theta_i, k) - p,$$

where k is the total number of drivers, including i, traveling and p is a transfer payment to the mechanism designer. We assume that for all θ_i and $k \leqslant n$, $v_{\theta_i} > 0$, that is $v(\cdot, \cdot)$ increases in θ. This implies that drivers with a higher VOT value the trip from O to D more highly. We further assume that there is congestion in the sense that more other drivers using the road reduce the value a driver obtains from the trip, i.e. for all θ_i and $k \leqslant n$, $v(\theta_i, k+1) - v(\theta_i, k) \leqslant 0$. Additionally, to interpret higher values of θ_i as a greater VOT, we assume that for drivers with a higher θ_i the impact of congestion is greater. This means that for all θ_i and $k \leqslant n$, $v_{\theta_i}(\theta_i, k+1) - v_{\theta_i}(\theta_i, k) \leqslant 0$. Finally we assume for all θ_i, $k \leqslant n$ and $k' < k$, $v(\theta_i, k+1) - v(\theta_i, k) \leqslant v(\theta_i, k'+1) - v(\theta_i, k')$. This means that the increase in travel time caused by congestion increases in the level of congestion. A sufficient condition is that travel time is convex in the number of drivers, like in the widely used BPR cost function.[13]

An underlying assumption of the congestion function is that drivers do not care about the identity of other drivers on the road, but only about their number. We believe this is a natural assumption in the context of transportation, and it simplifies the externality problem significantly.

To make the problem interesting we assume that for all θ_i, $v(\theta_i, 1) > 0$ and $v(\theta_i, n) < 0$ so that it is always optimal to have a positive number of drivers, strictly less than n, traveling. The allocation of agent i is given by $x_i \in \{0, 1\}$, where $x_i = 1$ means that i is allocated to traveling, while $x_i = 0$ means that i gets the outside option. We let $x \in X = \{0, 1\}^n$ denote the overall allocation. The number of traveling drivers is therefore given by $\sum_i^n x_i$. Thus, we can write $u_i\left(\theta_i, \sum_{j=1}^n x_j, p_i\right) = x_i u_i^T\left(\theta_i, \sum_{j=1}^n x_j, p_i\right) + (1 - x_i) u_i^O\left(\theta_i, \sum_{j=1}^n x_j, p_i\right)$.

[13] Congestion pricing is most relevant for medium levels of congestion where traffic is still flowing but at a lower speed. The model captures these situations. For larger levels of congestion, traffic comes to a standstill. While these situations are clearly important, a congestion-pricing scheme cannot resolve them, but only prevent them. For examples of papers employing convex cost function in the analysis of congestion, see Smith and Van Vuren (1993) or Mtoi and Moses (2014).

2.1 Example and Discussion of Modeling Assumptions

Ultimately drivers should only care about their travel time, but not about the number of other drivers. Because of congestion effects, the number of drivers on the road directly affects the travel time on that road. Therefore we can derive the reduced-form utility function specified above from an underlying model in which drivers only care about travel time. Too see this, suppose that $\tilde{v}(\theta_i, t)$ captures the utility to driver i from traveling from O to D in t units of time. If the time it takes to travel is given by $C(k)$ then we can derive $v(\theta_i, k)$ as $\tilde{v}(\theta_i, C(k))$.[14] Since more drivers on the road imply a longer travel time, we will often refer to travel time when discussing congestion.

We assume a quite general form of the utility function, but this specification encompasses a functional form commonly used in the literature.

$$v(\theta_i, k) = \theta_i(S - C(k)).$$

The travel time that captures congestion is given by $C(k)$. It is increasing and convex in the number of drivers on the road k and determines travel time. These assumptions are consistent, for example, with the BPR cost function.[15] The time available to the driver who makes the trip in the absence of congestion is S. We can then interpret $S - C(k)$ as the time not spent in traffic given the driver makes the trip, and $\theta_i S$ is the free-flow utility. How much a driver values time not spent in traffic is given by the VOT θ_i. The same functional form is used for example in Mayet and Hansen (2000).

The utility function $v(\theta_i, k)$ posits a direct relationship between VOT θ_i and the value for the trip $v(\theta_i, k)$. By assuming $v_{\theta_i} > 0$, we require that VOT and value for the trip are positively correlated.[16]

[14]Our assumptions on $v(\theta_i, k)$ can then be obtained by assuming that $\tilde{v}_t < 0$ and $\tilde{v}_{\theta_i} > 0$. We can obtain for all $k' < k$, $v(\theta_i, k+1) - v(\theta_i, k) \leqslant v(\theta_i, k'+1) - v(\theta_i, k')$ if $C(\cdot)$ is an increasing function with increasing differences and the derivative of $\tilde{v}_t$ is small enough. Furthermore the cross-partial derivative of $\tilde{v}(\cdot, \cdot)$ needs to be strictly negative to ensure that $v_{\theta_i}(\theta_i, k+1) - v_{\theta_i}(\theta_i, k) \leqslant 0$. This means that for drivers with a higher VOT θ_i, increases in travel time reduce utility by a greater amount.

[15]The BPR cost function is a fourth order polynomial of traffic volume on a road, see Smith and Van Vuren (1993) or Mtoi and Moses (2014) for examples of its application.

[16]Note that we allow $v_{\theta_i}(\theta_i, k)$ to depend on θ_i. Since one can interpret $v(\theta_i, k)$ as the average value for a trip for drivers with VOT θ_i, this specification can represent different degrees of positive correlations between VOT and value for a trip that depend on the level of θ_i. More generally, we only need to assume that the value function $v(\cdot)$ is monotone. This assumption is standard in the classic price-discrimination literature. Thus, we could also assume that $v(\cdot)$ decreases in θ_i, and we get the same qualitative results. If $v(\cdot)$ was non-monotone, i.e. increasing for some θ_i and

We think this assumption captures traffic situations where congestion is often a problem. For example in the context of morning commutes, a driver's value for the trip is highly correlated with the hourly wage she earns at work, and while in traffic she cannot earn this wage. Small (2012) discusses this morning-commuting problem and shows how utility maximization subject to time budget and budget constraints leads to a direct positive relationship between valuation of travel time and valuation for the trip.[17]

2.2 Mechanism Designer and First-Best Allocation

The mechanism designer maximizes total welfare, i.e. the sum of all drivers' utility plus the total revenue collected. Hence the objective function of the mechanism designer is given by:

$$W = max_{[x,p]} \sum_{i=1}^{n} \left[u_i \left(\theta_i, \sum_{j=1}^{n} x_j, p_i \right) + p_i \right] \tag{1}$$

In the absence of a mechanism, drivers would be free to travel or not without payments. In that case the value of the marginal trip would be close to zero.[18] Otherwise, more drivers would start to travel, thereby increasing travel time on this road. But there is also a coordination problem. Similar values for traveling and not traveling imply that there is no sorting into traveling or not according to the VOT. In contrast, the efficient allocation both solves the coordination problem, as drivers with a high VOT are allocated to travel, and ensures that traveling is strictly better to not traveling.

Before discussing the incentive problem that arises because each driver's VOT is private knowledge, we will analyze the mechanism designer's problem with perfect information.

We can cancel out the transfers and write the mechanism designer's problem as follows:

$$W = max_{[x]} \sum_{i=1}^{n} \left[x_i v \left(\theta_i, \sum_{j=1}^{n} x_j \right) \right] \tag{2}$$

decreasing for others, the constrained-optimal mechanism would involve pooling of some drivers with different VOTs.

[17] This formulation can allow for the VOT to depend on many other factors such as trip purpose, income, gender, and many more.

[18] This follows directly from Wardrop's first principle. For more on this see chapter 3.4.4 of Small and Verhoef (2007).

First, at an optimum the benefit of traveling needs to be positive. Otherwise, welfare could be increased by having just one traveling driver. Second, at the optimum drivers will be sorted according to their VOT, with high-value drivers traveling, while low value users do not travel. Suppose not, so that there is a non-traveling driver with a higher VOT than a traveling driver. Switching their allocations leaves the number of traveling drivers unchanged, but since $v_\theta > 0$, total welfare increases.

Lemma 1. *Let $x^{FB}(\theta)$ be the allocation that solves Equation 2 for a given θ. Then it must satisfy:*

- *For all θ_i, $v(\theta_i, \sum_{j=1}^n x_j^{FB}(\theta)) > 0$.*
- *If $\theta_i > \theta_j$, then $x_i^{FB}(\theta) \geqslant x_j^{FB}(\theta)$.*

Proof. For the first part suppose the utility from traveling is weakly negative under x^{FB}. Then consider letting only a single driver i use the road. By assumption for that driver we have $v(\theta_i, 1) > 0$, while all other drivers obtain zero. Hence welfare increased, which contradicts the optimality of x^{FB}.

For the second part suppose otherwise. Then there is at least one pair of drivers $\{i, j\}$ such that $\theta_i > \theta_j$ and $x_i^{FB}(\theta) < x_j^{FB}(\theta)$. But then one could replace $x_i^{FB}(\theta)$ and $x_j^{FB}(\theta)$ by $x_i^*(\theta) := x_j(\theta)$ and $x_j^*(\theta) := x_i(\theta)$ which would lead to change in welfare of

$$v\left(\theta_i, \sum_{j=1}^n x^{FB}(\theta)\right) - v\left(\theta_j, \sum_{j=1}^n x_j^{FB}(\theta)\right) = \int_{\theta_j}^{\theta_i} v_{\tilde{\theta}}\left(\tilde{\theta}, \sum_{j=1}^n x_j^{FB}(\theta)\right) d\tilde{\theta} > 0.$$

Hence welfare increased, which contradicts the optimality of x^{FB}. $\square$

Lemma 1 says that the first-best allocation has a simple structure: all drivers with a value of θ sufficiently high will travel while the remainder will not. To find the efficient allocation, we only need to find the optimal number of drivers using the road as a function of θ. We define $\theta^{(k)}$ to be the k^{th} highest value of θ from among the n drivers. Suppose there are $k \in \{0, 1, ..., n-1\}$ drivers on the road and consider adding another driver to it. The change in welfare resulting from this reallocation is given by:

$$\Delta_k(\theta) \equiv v(\theta^{(k+1)}, k+1) + \sum_{i=1}^{k} \big(v(\theta^{(i)}, k+1) - v(\theta^{(i)}, k) \big)$$

The expression Δ_k characterizes the key trade-off in determining the efficient allocation. The first term appearing in Δ_k is the benefit of allocating the driver with the $(k+1)$-highest VOT to traveling. The value of the trip for this driver is given by $v(\theta^{(k+1)}, k+1)$. The second term captures the cost of increased congestion for the first k drivers from another driver on the road. For efficient allocations the first term will generally be positive. The second term always enters negatively due to the assumption that the road is congestible, i.e. $v(\cdot, \cdot)$ is decreasing in its second argument. The next Lemma characterizes some useful properties of $\Delta_k(\theta)$ that we will use in the proof of Proposition 1.

Lemma 2. *For all $\theta \in \Theta$ and all $k \in \{0, 1, ..., n-2\}$, we have that $\Delta_k(\theta) > \Delta_{k+1}(\theta)$. Furthermore $\Delta_0(\theta) > 0$.*

Proof. For the first result consider the difference in the first terms of $\Delta_k(\theta)$ and $\Delta_{k+1}(\theta)$ which is given by:

$$\begin{aligned} & v(\theta^{(k+2)}, k+2) - v(\theta^{(k+1)}, k+1) \\ + \sum_{i=1}^{k} \big[\big(v(\theta^{(i)}, k+2) - v(\theta^{(i)}, k+1) \big) - & \big(v(\theta^{(i)}, k+1) - v(\theta^{(i)}, k) \big) \big] \\ & + v(\theta^{(k+1)}, k+2) - v(\theta^{(k+1)}, k+1) \end{aligned}$$

From the definition of $\theta^{(k)}$ we have that $\theta^{(k)} > \theta^{(k+1)}$. Furthermore we have that $v(\cdot, \cdot)$ is strictly increasing in its first argument and decreasing in its second argument, so that the first difference is strictly negative. Next consider the second term which is given by:

$$\sum_{i=1}^{k} \big[\big(v(\theta^{(i)}, k+2) - v(\theta^{(i)}, k+1) \big) - \big(v(\theta^{(i)}, k+1) - v(\theta^{(i)}, k) \big) \big]$$

This term is weakly negative, which follows from our assumption that for all θ_i, k and $k' < k$, $v(\theta_i, k+1) - v(\theta_i, k) \leqslant v(\theta_i, k'+1) - v(\theta_i, k')$. Finally consider the third term. This term is also weakly negative, which follows from $v(\cdot, k)$ being decreasing in k. Hence it follows that

$\Delta_k(\theta) > \Delta_{k+1}(\theta)$.

We can verify the value of $\Delta_0(\theta)$ from the definition of $\Delta_k(\theta)$ by setting $k = 0$. $\square$

Lemma 2 shows how the trade-off from adding another driver evolves as more drivers travel. First, the benefit of adding one driver to the traffic falls, since the VOT of the marginal driver decreases. Second, the costs which an additional driver imposes on the other users of the road increase, because more drivers are affected by increased congestion. Both effects go in the same direction, so that the welfare gain of each additional driver falls. When there is no driver on the road, then there is no cost of adding the first driver. The following Proposition follows directly from Lemma 2 and gives a necessary and sufficient condition for the first-best allocation.

Proposition 1. *There exists some $k^* < n$ such that for all $i = \{1, ..., n\}$ at the optimum $x_i(\theta) = 1$ if and only if $\theta_i \geqslant \theta^{(k^*)}$. The value of k^* is given by:*

$$k^* = \max_{\Delta_k(\theta) \geqslant 0} k + 1 \tag{3}$$

The Proposition characterizes the efficient number of drivers k^*. We will often refer to x^{FB} as the efficient allocation that leads to the efficient number of drivers k^*. The logic behind Proposition 1 follows from Lemma 2: as the number of traveling drivers increases, the benefit of adding another driver shrinks. Welfare changes are strictly positive when the driver with the largest valuation is the only driver on the road, but become smaller as more drivers travel. By assumption there is an interior solution, i.e. $v(\cdot, n) < 0$ implies that $k^* < n$.

In the next Lemma, we derive comparative statics of the efficient allocation with respect to θ, the vector of the drivers' VOT.

Lemma 3.

1. *Comparative statics of the efficient number of drivers k^*:*

   (a) *If $\theta_i' \geqslant \theta_l(\theta_{-i})$, then $k^*(\theta_i, \theta_{-i})$ is a weakly decreasing function in θ_i.*

   (b) *If $\theta_i' < \theta_l(\theta_{-i})$, then $k^*(\theta_i, \theta_{-i})$ is constant for $\theta_i < \theta'_{-i}$, increases by one at $\theta_i = \theta_i'$ and for $\theta_i > \theta'_{-i}$ is weakly increasing.*

2. *The utility $x_i^{FB} v(\theta_i, \sum_{j=1}^n x_j^{FB})$ of driver i at the efficient allocation is weakly decreasing in θ_i, $\forall i$.*

Proof. For the results we need the effect of a change in a single θ_i on the optimal allocation, holding θ_{-i} fixed. Let $\theta_{-i}^{(k)}$ be the k^{th} highest VOT among all drivers except driver i. Consider the auxiliary problem in which we set $\theta_i = 0$, but driver i still needs to be allocated. In that case, it is clear that $x_i = 0$ is optimal. Denote the efficient allocation in this case by k_{-i}^*, which is a function of θ_{-i}. Refer to the driver with the k_{-i}^* highest VOT as driver $l(\theta_{-i})$. Denote his associated VOT by $\theta_l(\theta_{-i}) = \theta_{-i}^{k_{-i}^*}$. We denote by θ_i' the value of θ_i such that adding driver i to the road when k_{-i}^* are allocated to it results in no welfare change. Thus, θ_i' is defined by:

$$v(\theta_i', k_{-i}^* + 1) - \sum_{j=1}^{k_{-i}^*} \left[v(\theta^{(j)}, k_{-i}^* + 1) - v(\theta^{(j)}, k_{-i}^*) \right] = 0$$

We consider two cases. First, suppose $\theta_i' \geqslant \theta_l(\theta_{-i})$. Then for $\theta_i \leqslant \theta_l(\theta_{-i})$ we have that k^* does not vary in θ_i and neither does the allocation x^{FB}. When $\theta_i > \theta_l(\theta_{-i})$ it follows that $x_i^{FB} = 1$ while $x_l^{FB} = 0$. As θ_i increases, k^* falls. In this case, θ_i is allocated to traveling, and the cost of adding other drivers increases as the VOT of driver i increases. Therefore k^* is a decreasing function of θ_i.

Second, suppose $\theta_i' < \theta_l(\theta_{-i})$. For $\theta_i < \theta_i'$ we have $k^* = k_{-i}^*$. For $\theta_i \in [\theta_i', \theta_l(\theta_{-i})]$ it is optimal to add driver i to the road without removing any other driver from it. Therefore we have $k^* = k_{-i}^* + 1$. For $\theta_i > \theta_l(\theta_{-i})$, k^* falls as θ_i increases. Since θ_i is allocated to the road, any increase in θ_i increases the cost of adding other drivers to it, implying that k^* will fall. $\square$

Intuitively, if θ_i increases and becomes just large enough to be assigned to traveling, driver i will be either added to the set of traveling drivers or she replaces someone else. At the point when driver i is added to the road, k^* can only be non-decreasing in θ_i. This implies that the number of traveling drivers is only non-decreasing in θ_i at that point. Nonetheless, the utility at the efficient allocation for driver i is strictly larger due to Lemma 1, since for a lower θ_i that particular driver is assigned to the outside option. When driver i is assigned to traveling and θ_i increases further, k^* decreases and the utility obtained by driver i increases.

3 Implementing the First-Best Allocation

3.1 Theoretical Analysis

In this section we consider the problem of allocating drivers to traveling or not traveling as a mechanism design problem. We derive an efficient and incentive-compatible mechanism.

An important consideration in the mechanism design literature is incentive compatibility, i.e. designing mechanisms in a way that drivers always reveal their private information truthfully.[19] As a first step, we apply the dominant-strategy *revelation principle* of Gibbard (1973), which allows us to study a large class of mechanisms by focusing on a smaller subclass. By the revelation principle every complicated mechanism involving potentially very large message spaces can be replaced by a simpler mechanism that only asks drivers to directly report their type truthfully. Thus, instead of studying complicated mechanisms where in equilibrium a type can be inferred from a message, it is without loss of generality to study a direct mechanism. In such a mechanism, each driver will be asked to report her private information, namely her valuation for travel time. The mechanism designer then allocates drivers to traveling or outside option based on this information. More precisely, a direct mechanism is a function associating to each θ an allocation, x and a transfer function p, where p is the transfer paid by a driver to the mechanism designer. In short, a mechanism is a mapping from reports of the drivers' valuation of time to an allocation and transfers, i.e. $[x(\theta), p(\theta)] : \theta \to X \times \mathbb{R}^n$.

More intuitively, the revelation principle allows us to translate a problem where drivers reveal their preferences by choosing from a menu of prices into a problem where drivers report their private information to the mechanism, and the mechanism commits to make the optimal choice for the drivers. When we ue the term 'report' private information, we want to emphasize that we talk about the direct mechanism.

We apply the concept of dominant-strategy incentive compatibility. The mechanism designer requires each driver prefers to truthfully reveal her VOT given all possible valuations of the other

[19]In our context, truthfully revealing private information means that each driver actually chooses the option which the mechanism designer wants them to take.

drivers. Hence the mechanism works regardless of what driver i believes about the distribution of driver j's valuation of travel time and the mechanism designer obtains the exact valuations of the n drivers without requiring precise information ex-ante.

Definition 1. *A direct mechanism $[x, p]$ is dominant-strategy incentive compatible if for all $\theta_i, \hat{\theta}_i \in \Theta_i$ and $\theta_{-i} \in \Theta_{-i}$ it satisfies*

$$U_i(\theta_i;\theta_{-i}) \equiv x_i(\theta)v(\theta_i, \sum_{j=1}^{n} x_j(\theta)) - p_i(\theta) \geqslant x_i(\hat{\theta}_i, \theta_{-i})v(\theta_i, \sum_{j=1}^{n} x_j(\hat{\theta}_i, \theta_{-i})) - p_i(\hat{\theta}_i, \theta_{-i})$$
$$\equiv U_i(\hat{\theta}_i, \theta_i; \theta_{-i}) \qquad \text{(DIC)}$$

A mechanism that satisfies DIC offers attractive features in our application. For a driver with type θ_i it is optimal to reveal this value truthfully, rather than any other value $\hat{\theta}_i$. This is true for all other possible revealed values θ_{-i} of the other drivers. Requiring that truth-telling is optimal for all possible realizations of other drivers' values θ_{-i} implies that truth-telling is optimal for each driver i no matter what her beliefs are about other drivers' valuations.[20] This means that drivers want to reveal their travel preferences truthfully, whether they have correct beliefs about other drivers' travel preferences or not. Hence, mechanisms that satisfy DIC do not require detailed knowledge of drivers about other drivers' VOTs. Also the regulator does not need detailed knowledge about the distribution of the drivers' VOTs, making these mechanisms also robust to incorrect beliefs of the regulator.

The mechanism designer maximizes welfare given by equation 1 subject to the DIC constraints. We say that an allocation function $x(\theta)$ is *implemented in dominant strategies* by payment rules $p_i(\theta)$ if they satisfy the incentive-compatibility constraints. Since the mechanism designer maximizes total welfare, there is no revenue-raising motive. Note that the private information held by the drivers affects other drivers only indirectly through the resulting allocation.

In the following discussion, we make use of some extra notation. We denote by $\theta^{(k^*)}$ the k^*-highest VOT, where k^* is as defined in Proposition 1. We suppress the dependence of k^* on θ for

[20] Alternatively one could consider the incentives of drivers to report truthfully given their beliefs about the types of other drivers. However in our application it is unlikely that the mechanism designer knows these beliefs.

simplicity. Let $k^*_{-i}(\theta_{-i})$ be the value of k^* as in Proposition 1 excluding driver i. As before we let $\theta^{(k)}_{-i}$ be the k^{th}-highest VOT in the problem excluding driver i among the $n-1$ remaining drivers. Again, we suppress the dependence of k^*_{-i} on θ_{-i} for simplicity. For notational ease we define for each $i \in \{1, ..., n\}$, the following two sets:

$$\Omega_i^+(\theta) \equiv \{j \neq i | \theta_j \geqslant \theta^{(k^*)}\}$$
$$\Omega_i^0(\theta) \equiv \left\{j \neq i | \theta_j \in \left(\theta_l(\theta_{-i}), \theta^{(k^*)}\right)\right\}$$

Ω_i^+ is the set of drivers, excluding i, that travel irrespective of driver i's allocation. These drivers are affected by driver i only through changes in congestion. Similarly,Ω_i^0 is the set of drivers assigned to travel if and only if driver i is allocated to the outside option. These drivers are affected by driver i through the impact of his allocation on theirs. If the set Ω_i^0 is empty, then driver i's report does not affect the assignment of other drivers to the road, but may affect the level of congestion on it. Note that both Ω_i^+ and Ω_i^0 depend on the vector of the drivers' VOT, θ. Note also that a consequence of Lemma 3 is that $k^* - k^*_{-i} \leqslant 1$.

Proposition 2. *The following payment rule implements the first-best allocation, $x^{FB}(\theta)$ and specifies for all $i \in \{1, ..., n\}$:*

$$p_i^{FB}(\theta) = -\sum_{j \in \Omega_i^+} \left[v(\theta_j, k^*) - v(\theta_j, k^*_{-i})\right] + \sum_{j \in \Omega_i^0} v(\theta_j, k^*_{-i}) \tag{4}$$

Proof. The payment rule is an application of the Vickrey-Clarke-Groves mechanism. The basic logic behind such a mechanism is to make each driver residual claimant of total welfare and thereby let them internalize the mechanism designer's problem. The incentives for truth-telling are unaffected by an added term which does not depend on a driver's own report, but it may depend on the reports of all the other drivers. In our model, this implies that the VCG payment rule has the following form:

$$p_i^{VCG}(\hat{\theta}_i, \theta_{-i}) = -\sum_{j=1; j \neq i}^{n} x_j(\hat{\theta}_j, \theta_{-i}) v(\theta_j, k^*(\hat{\theta}_i, \theta_{-i})) + h_i(\theta_{-i})$$

Since $k^*(\theta)$ is by definition the function that maximizes welfare for each $\theta \in \Theta$, driver i maximizes her utility, given by:

$$U_i(\hat{\theta}_i, \theta_i; \theta_{-i}) = x_i(\hat{\theta}_i, \theta_{-i}) v(\theta_i, k^*(\hat{\theta}_i, \theta_{-i})) \\ + \sum_{j=1; j \neq i}^{n} x_j(\hat{\theta}_i, \theta_{-i}) v(\theta_j, k*) - h_i(\theta_{-i})$$

This means that driver i faces the mechanism designer's problem 1, so that reporting $\hat{\theta}_i = \theta_i$ is optimal. The function $h_i(\theta_{-i})$ is constant in $\hat{\theta}_i$ and therefore does not affect i's incentives. We choose $h_i(\theta_{-i})$ as the sum of utilities of all drivers but i in the hypothetical scenario in which driver i was excluded. This induce the desirable property that the outside option is for free. We denote by $x^{FB}_{j,-i}$ the optimal allocation of driver j when driver i is excluded.

$$h_i(\theta_{-i}) \equiv \sum_{j=1; j \neq i}^{n} x^{FB}_{j,-i}(\theta_{-i}) v(\theta_j, k^*_{-i})$$

Since $k^*_{-i}(\theta_{-i})$ solves the allocation problem as if the VOT of driver i was zero, driver i will then be allocated to the outside option. Hence the payment above is the surplus of the other drivers under the optimal allocation given that i is assigned to the outside option. Note that it may happen that $\theta_{-i}^{(k^*_{-i})} > \theta^{(k^*)}$. In this case, the set Ω^0_i is empty and only the first term in 4 remains. Substituting our choice of $h_i(\theta_{-i})$ into p^{VCG}_i and applying the definitions of Ω^+_i and Ω^0_i leads to 4. □

The payment schedule 4 consists of two components. We call the first term the congestion effect: $-\sum_{j \in \Omega^+_i} \left[v(\theta_j, k^*) - v(\theta_j, k^*_{-i}) \right]$. It captures how driver i's report changes the travel time on the road. This effect may be either positive or negative. It will be negative (meaning that driver i has to pay less) under the condition in Part 1(a) of Lemma 3. In that case a higher report of θ_i reduces k^*, the number of drivers on the road. All drivers remaining on the road will benefit from reduced congestion. The reduction in the price paid by driver i reflects the value of this reduced congestion of the other drivers. The first component will be positive (meaning driver i has to pay more) only if the valuations of time θ imply $k^* - k^*_{-i} = 1$.[21] In that case the report of driver i

[21] This corresponds to case 1(b) in Lemma 3 at the point where k^* is increasing in θ_i. Intuitively, this happens when a driver i's report causes no other driver to be allocated from traveling to the outside option.

increases the number of traveling drivers and thereby congestion. When $k^* = k^*_{-i}$, the congestion effect is zero. In that case driver i replaces another driver on the road.

The second component is the reallocation effect: $\sum_{j \in \Omega_i^0} v(\theta_j, k^*_{-i})$. It captures how the report of driver i induces a reallocation of other drivers from the road to the outside option. For high values of θ_i, it becomes efficient to reduce congestion. This is accomplished by reallocating drivers with a lower VOT to the outside option. It follows from Lemma 1 that those who are reallocated obtain lower utility. The reallocation effect reflects these costs. In contrast, drivers allocated to the outside option pay nothing. By being allocated to the outside option, both the congestion and the reallocation effects are zero.[22]

Remark 1. *The payment schedule $p_i^{FB}(\theta)$ is a weakly increasing step function.*

The payment schedule 4 depends on θ_i only through its effect on the efficient allocation $k^*(\theta)$. By definition, k^* can only take a finite number of values. It is constant almost everywhere but there are jumps at a finite number of points, whenever the number of drivers allocated to traveling changes. Therefore the payment schedule faced by driver i is constant almost everywhere and has a finite number of jumps. Furthermore by Lemma 3 the utility of driver i is weakly increasing in θ_i. This implies that the payment schedule of driver i is weakly increasing. If it were not, there would be cases in which driver i could misreport her valuation to obtain both a lower travel time and a lower payment. This would violate incentive compatibility. Most importantly, a single Pigouvian price does not implement the efficient allocation and is therefore not generally optimal. The following example illustrates the step function of the price schedule.

3.2 Examples

3.2.1 Example 1: Illustration of the Price Schedule.

In order to illustrate the pricing schedule, consider a problem in which only the VOT of one driver is unknown. There are five drivers $i = 0, 1, 2, 3, 4$. Driver i's utility from traveling is $U_i^T =$

[22] Note that if the outside option is another congestible road, prices for using this congestible alternative may not be zero. We discuss this in the supplementary online material 6.1.

$v(\theta_i, k) - p = \theta_i(2 - 0.5k) - p.$

For this particular example we derive the utility of driver i from the following fundamentals. Reaching the destination without travel time gives driver i a net utility $\theta_i 2.5$. It takes a driver half a minute to travel to her destination if there are no other drivers. For each driver on the road, the average travel time for all drivers on that road increases by half a minute, that is travel time follows the following linear function $C(k) = 0.5 + 0.5k$. Taken together, these fundamentals induce the aforementioned utility of traveling.

It is common knowledge that the VOT is given by $\theta_i = 11 - i$ for $i \in \{1, 2, 3, 4\}$. The valuation of time of driver $i = 0$, given by $\theta_0 \in (0, \infty)$ is unobservable private information. Let $\mathcal{A}^*(\theta_0)$ be the set of drivers who travel at the efficient allocation.

The efficient allocation depends on the value of θ_0 which we refer to as aggregate uncertainty. Knowing θ_0 the mechanism designer could set a single price as a function of θ_0 such that drivers use roads efficiently. When the mechanism designer does not know θ_0, a mechanism which sets the price as a function of the report of θ_0 gives driver 0 an incentive to lie about her VOT. However the efficient allocation can be implemented by letting driver 0 face the payment schedule $P^*(t)$ for a travel time $t \in \{0.5, 1, 1.5, 2.5\}$.

$$P^*(t) = \begin{cases} 0 & \text{outside option} \\ 9.5 & t = 1.5 \\ 14.4 & t = 1 \\ 30.4 & t = 0.5 \end{cases} \qquad \mathcal{A}^*(\theta_0) = \begin{cases} \{1,2\} & \theta_0 < 8.\overline{63} \\ \{0,1,2\} & 8.\overline{63} \leqslant \theta_0 < 9.8 \\ \{0,1\} & 9.8 \leqslant \theta_0 < 32 \\ \{0\} & 32 \leqslant \theta_0 \end{cases}$$

The difference of this payment schedule to setting a single price is that it allows to charge driver 0 different prices for different travel times, while a single price mechanism only charges for use of the road irrespective of the number of other drivers on the road. This payment schedule is constructed in a way that the prices faced by driver 0 capture the externalities this driver imposes on the other drivers and gives driver 0 incentives to report θ_0 truthfully. Figure 1 displays the resulting price schedule and the efficient allocation as a function of driver 0's value of travel time.

In the previous example, only one driver has private information about its VOT. We now discuss an example where the regulator does not know the VOT of multiple drivers.

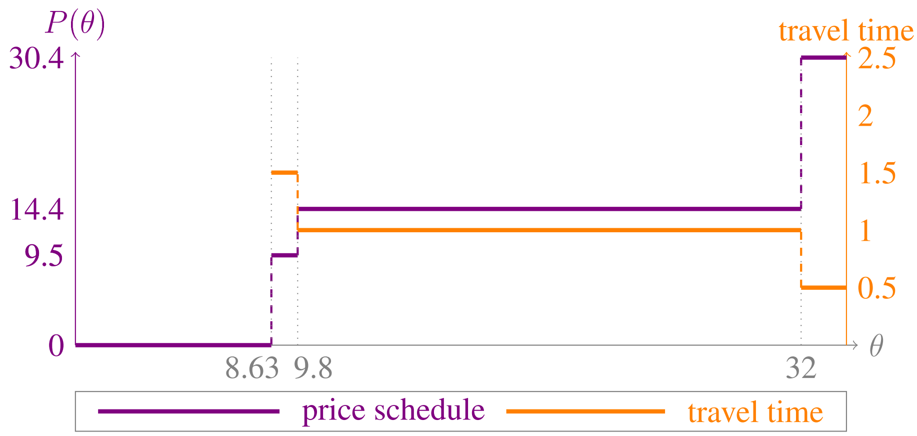

Figure 1: Efficient Allocation and Pricing

3.2.2 Example 2: Two Drivers

To illustrate, consider an example with two drivers. Travel utility is $v(\theta_i, k) = \theta_i(4-k)$, while utility from the outside option is constant at zero. This could be derived from fundamentals as follows. The utility of traveling is given by $\theta_i(4-t)$, where $\theta_i 4$ is the benefit from reaching D with minimal travel time and t is the increase in travel time due to congestion. If we assume that travel time on the road is given by $t = C(k) = k$, we can derive the utility function above.

If $\theta_i > \theta_j$ it is efficient for both drivers to travel if and only if $-2(\theta_i + \theta_j) > -\theta_i - 4\theta_j$. This simplifies to $\theta_i < 2\theta_j$. Otherwise only θ_i is efficiently allocated to traveling. The efficient allocation is

$$(x_1, x_2)(\theta) = \begin{cases} \{1,1\} & if \quad \theta_1 \in [\frac{1}{2}\theta_2, 2\theta_2] \\ \{1,0\} & if \quad \theta_1 > 2\theta_2 \\ \{0,1\} & if \quad \theta_1 < \frac{1}{2}\theta_2 \end{cases}$$

Figure 2 shows that when the VOTs are similar (i.e. we are in the violet area around the 45 degree line), it is efficient for both of them to travel. When the relative difference in the VOTs is large, it is optimal to assign only the driver with the higher VOT to the road.

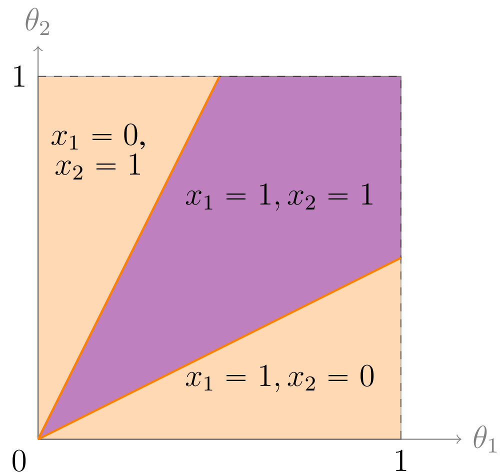

Figure 2: Optimal allocations with Two Drivers

We apply Proposition 2 to get the price schedule that implements this allocation for driver 1, together with the according travel times:

$$p_1(\theta) = \begin{cases} 0 & if \quad \theta_1 < \frac{1}{2}\theta_2 \\ \theta_2 & if \quad \theta_1 \in [\frac{1}{2}\theta_2, 2\theta_2] \\ 3\theta_2 & if \quad \theta_1 > 2\theta_2 \end{cases} \quad t_1(\theta) = \begin{cases} \textbf{no travel} & if \quad \theta_1 < \frac{1}{2}\theta_2 \\ 2 & if \quad \theta_1 \in [\frac{1}{2}\theta_2, 2\theta_2] \\ 1 & if \quad \theta_1 > 2\theta_2 \end{cases} \tag{5}$$

These schedules also illustrate that we can express the price schedule of driver 1 $p_1(\theta)$ also in terms of travel time $t_1(\theta)$, leading to a schedule of prices per travel time.

Figure 3 (left) plots the optimal price schedule faced by driver 1 for two values of θ_2. Note that the price paid by driver 1 is not monotone in θ_2.

Figure 3 (right) plots the optimal travel time of driver 1 for two different values of θ_2 as a function of θ_1. Unlike the payment schedule, the travel time is a monotone function of the value of θ_2. The payment schedule is independent of the distribution of θ.

We use these examples to illustrate the implementation of the proposed pricing mechanism in practice. One can think of the mechanism as following a two-stage procedure. In the first stage drivers report their travel preferences. They could report their VOT directly. To simplify, drivers

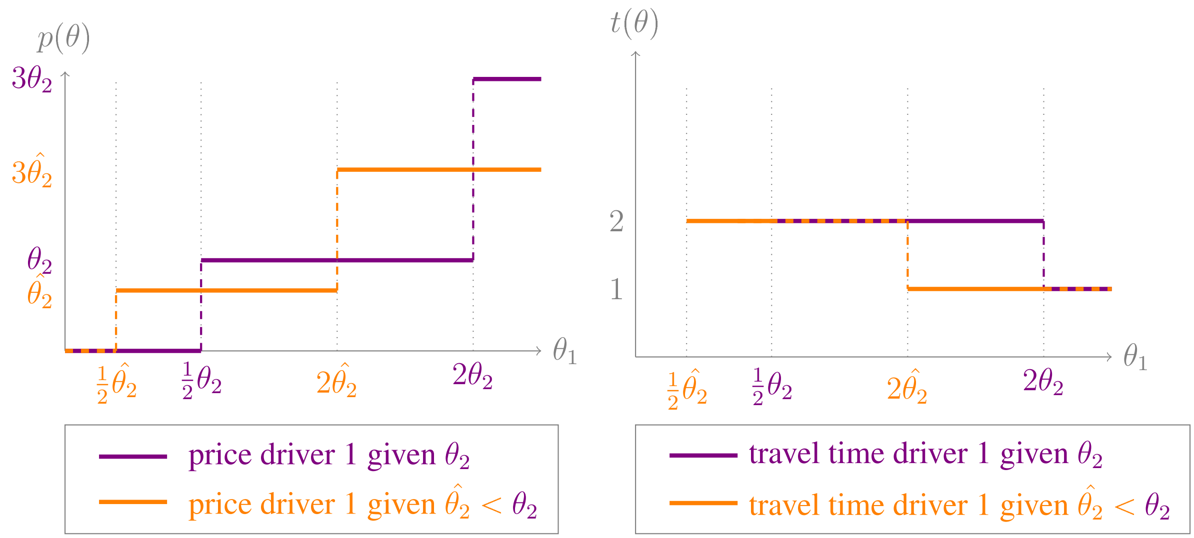

Figure 3: Payment schedule (left) and travel time (right) for two different values of θ_2

could choose between some classifications of urgency. In the second stage the mechanism uses these reports to build the pricing schedule that we derived above and drivers can choose whether to travel or not. Since the mechanism knows from their initial report what is best for each driver, the mechanism can also automatically choose for drivers at the second stage.

The practical implementation of the pricing scheme that is closest to the direct mechanism discussed above would be similar to a second-price auction. Drivers who want to use the road at a given departure time report their VOT, i.e. their willingness to pay to arrive a unit of time earlier. Just before departure time, the mechanism would have gathered all the VOTs of drivers who want to depart at the same time.[23] The mechanism could then construct the pricing schedules, for example $p_1(\theta)$ and $t_1(\theta)$ in the second example given above. Recall that the schedule of driver i does not depend on her own valuation. Misreporting her VOT therefore does not change her price schedule. At this stage the regulator knows what is best for the drivers if their initial report was truthful, hence the mechanism can allocate drivers automatically.The mechanism could also let drivers choose themselves at this stage with the requirement that their departure time might be delayed if their choice is inconsistent with earlier reports of the VOT.

[23] In practice, the mechanism could require consumers to make their reports until some time interval, e.g. 5 minutes, before departure. For practicability, the mechanism could prespecify regular departure times in a given interval, e.g. every 1/2/5 minutes.

From the driver's point of view, this procedure is similar to bidding on ebay where they can indicate their maximum willingness to pay for an object. Drivers only need to report their VOT, and the mechanism would allocate them automatically according to what is best for them given these report. They would never pay more that what they indicated as their VOT. Similar to a second-price auction, drivers that report a VOT that is higher than their true value might have to pay more than their true value. Drivers who underreport their VOT might travel slower than they would prefer.

Both proposed procedures require drivers to communicate with the regulator instantaneously. Given the wide-spread use of smart phones and GPS- or satellite navigation systems in cars, we do not believe this to be a major obstacle. Both mechanisms could be programmed as a smart-phone app, and the smart-phones GPS can be used to verify that the driver actually uses the intended road.

4 Comparison to Pigouvian Pricing

4.1 Congestion Pricing in the Limit

Our results so far suggest that each driver should face a price schedule that depends on other driver's revealed travel preferences. To reconcile with the Pigouvian approach we consider a variant of our model in which the number of drivers goes to infinity. To keep the analysis interesting, we assume that travel capacity also increases with the number of drivers in the economy. For simplicity we assume that $v(\theta_i, k) = \theta_i(S - b_n k)$, where $b_n = b/n$ and b, S are a positive constants. As in Section 2.1, $\theta_i \cdot S$ is the free-flow benefit of traveling for driver i. $b_n \cdot k$ is the travel time on the road. The marginal effect on congestion is $b_n = b/n$. The constant b measures the strength of congestion. The denominator n reflects that the capacity on the road increases with the number of drivers n in the economy.[24] As in the examples before, for a given number of drivers n, the marginal effect of

[24]We need to assume that road capacity increases with n to keep the analysis interesting. This assumption implies that as the number of drivers *in the economy* n increases, also road capacity increases and the effect of each driver on congestion becomes "small". Without this assumption, the share of drivers on the road in the limit would always be zero. In many settings this assumption is likely to be violated, but it is necessary to bridge the gap between the Pigouvian approach and our more general setup. Since in reality, and especially in cities, space is limited, traffic would

congestion of each driver is constant.

We assume $S - b < 0$, to ensure that at the efficient allocation not all drivers travel in the limit. We also normalize the objective function by dividing through by the number of drivers n:

$$W = max_{[x_i(\theta), p(\theta)]} \frac{1}{n} \sum_{i=1}^{n} [u_i(p_i, x; \theta_i) + p_i] \tag{6}$$

For each n we can use the results of Lemmas 1 and 2 as well as Proposition 1 to find the optimal solution. We consider now the probability limit of the value of $\Delta_k(\theta)$ as n converges to infinity, while letting k go to infinity, such that $\lim_{n\to\infty} \frac{k}{n} = q \in [0, 1]$.

$$\text{plim}_{n\to\infty} \Delta_k = \text{plim}_{n\to\infty} \theta^{(k+1)} \left(S - \frac{k+1}{n} b \right) - \frac{b}{n} \sum_{i=1}^{n} \theta_i \mathbb{1} \left(\theta_i > \theta^{(k+1)} \right) \tag{7}$$

Note that $\text{plim}_{n\to\infty} \theta^{(k+1)}$ is simply the $(1-q)^{th}$ quantile of the distribution function $F(\cdot)$, which we will denote by $\theta^q \equiv F^{-1}(1 - q)$. The probability limit of the second term is:

$$\text{plim}_{n\to\infty} \frac{b}{n} \sum_{i=1}^{n} \theta_i \mathbb{1} \left(\theta_i > \theta^{(k+1)} \right) = b \int_{\theta^q}^{\overline{\theta}} \theta dF(\theta)$$

Therefore, we have that:

$$\text{plim}_{n\to\infty} \Delta_k = \theta^q \left(S - qb \right) - b \int_{\theta^q}^{\overline{\theta}} \theta dF(\theta) \equiv \Delta_q \tag{8}$$

The welfare change Δ_q is strictly decreasing in q. We denote by q^* the efficient level of q, given by the unique solution of $\Delta_q = 0$. Let θ^* be such that a fraction q^* of all drivers that have a greater VOT than θ^*. The optimality condition in the limit is then given by:

$$\theta^* \left(S - q^* b \right) = b \int_{\theta^*}^{\overline{\theta}} \theta dF(\theta) \tag{9}$$

Consistent with the Pigouvian approach a deterministic share of drivers efficiently travels in the limit. Therefore, we can set a single congestion charge to implement the efficient allocation in the limit. The following Proposition summarizes the preceding discussion:

come to a stand still if the number of drivers gets too large.

Proposition 3. *In the limit, a unique Pigouvian price p^* for traveling implements the efficient allocation. This price is given by:*

$$p^* = b \int_{\theta*}^{\overline{\theta}} \theta dF(\theta) \tag{10}$$

The efficient congestion charge equals the value of the marginal increase in travel time from a small increase in the number of traveling drivers. The Pigouvian approach holds only if there are infinitely many small drivers. Intuitively, each driver has a negligible effect on the other drivers in the limit. To set the Pigouvian congestion charge correctly, the mechanism designer needs to know the distribution of the VOT. In contrast, the optimal mechanism of Proposition 2 does not require any prior knowledge of this distribution.

4.2 Simulations

The limit results of the previous section suggest that an appropriately set Pigouvian price maximizes welfare. There are two impediments to setting the Pigouvian price optimally. First, in practice the number of drivers is finite. Given that we focus on a static problem in which all drivers use the road simultaneously, it is likely that the number of drivers in applications is low. Especially in cities, traffic is possible only if the number of drivers is not too large since otherwise it turns into a standstill. In other traffic-related applications like shipping or air travel, the number of participants is frequently quite low as well. In these cases, one cannot rely on limit results and the more complex price schedule that we introduce in this paper is optimal.

Second, setting the Pigouvian price optimally requires knowledge of the distribution of the VOT. As this is usually not known by the policy maker, the price will be set either as a function of the policy maker's prior or may need to be estimated. But even when a policy maker can estimate the distribution of unobservable factors correctly, she cannot condition prices on the particular unobservable factors of each individual driver. An example of this would be a driver having an important meeting.

Both issues are no longer relevant when the mechanism designer asks drivers directly about their valuation of travel time and sets payment schedules that induce truthful revelation of these information.

To illustrate potential problems with mechanism designers' priors, suppose the common distribution of the VOT depends on an unobservable parameter α, so that each θ_i is distributed according to $F(\theta_i; \alpha)$. We assume that α is distributed uniformly over the unit interval. This implies that from the mechanism designer's view, the values of θ_i are not independent. The mechanism designer's prior is then $f_p(\theta_i) \equiv \int_0^1 f(\theta_i; \alpha) d\alpha$. The mechanism designer sets the Pigouvian price based on this prior as follows:

$$p = b \int_{\theta*}^{\bar{\theta}} \theta dF_p(\theta)$$

This will not yield the optimal price $p^*(\alpha)$ that the mechanism designer would set if he knew α. In contrast, our mechanism always achieves an efficient allocation. So far aggregate uncertainty resulted from a finite number of drivers. The mechanism designer's lack of knowledge over α is another source of aggregate uncertainty. Therefore, aggregate uncertainty implies that a Pigouvian price based on priors is not efficient even with a continuum of drivers.

We analyze the performance of Pigouvian prices using a simulation. We calculate the normalized welfare loss[25] resulting under the Pigouvian price. To capture implications of a regulator with wrong prior information on demand, we also consider welfare losses under pricing errors ($\pm 20\%$).

The simulation compares welfare under the efficient limit price (Pigouvian price) $\pm 20\%$ to maximum welfare in %. For each n, 10,000 random draws were taken. $v(\theta_i, k) = \theta_i(S - (b/n) \cdot k)$ as in the last Section, where $b = 15$ and $S = 14$. We choose a lognormal distribution for $F(\cdot)$ with a median of 21.46 \$/h and interquartile range of 10.47 \$/h, taken from Small et al. (2005), Table 3. Hence the natural logarithm of the VOT is distributed with a mean of 3.07 and a variance of 0.13. The efficient limit price under this set-up is given by 186\$, which implies that around 40.2% of drivers use the road in the limit.

[25]We calculate the welfare loss as follows: $Loss = \frac{W* - W(p)}{W* - W(0)}$, where $W*$ is maximum welfare, $W(p)$ gives welfare when a price of p is set and $W(0)$ is welfare resulting from a price of 0 and all drivers' utility equals their outside option. We compare the loss in welfare from setting a price p to the welfare loss from not having a mechanism. Thus welfare losses are normalized by the maximal welfare gain from a mechanism.

The results are summarized in Figure 4. The purple curve indicates the welfare loss for the correct Pigouvian limit price. This welfare loss vanishes as the number of drivers increases. But for a small number of drivers there is a welfare loss. The welfare loss when a price is set at 20% below the correct Pigouvian price is in orange. It is clear that the welfare loss under this price converges to a strictly positive value.

When there are few drivers the lower price gives a lower welfare loss than the correct Pigouvian price. This can result from realizations of the valuations of time such that none of the drivers is willing to pay the price for traveling. In that case a lower price may induce at least one driver to travel, which always dominates no driver traveling. If we restricted attention to a single price, the optimal price under this constraint depends on n. The blue curve shows the welfare loss when the price is set at 20% above the optimal Pigouvian price. Again, the welfare loss does not vanish as the number of drivers increases.

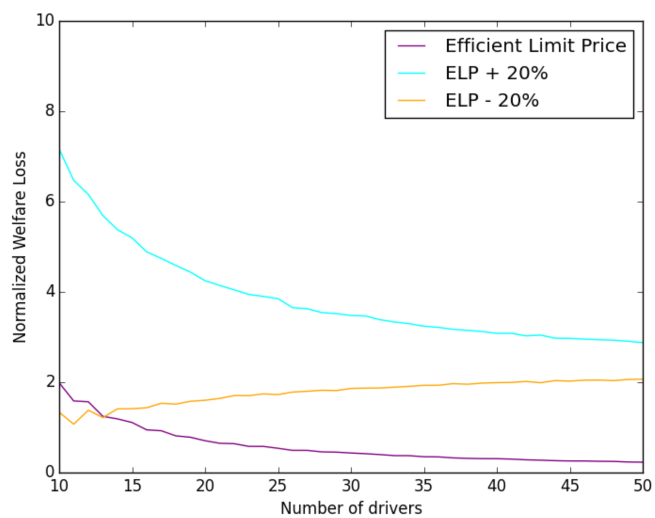

Figure 4: Welfare loss of single price mechanisms in the limit

The simulation highlights that even when there are infinitely many small drivers our mechanism can achieve significant welfare gains by acquiring the information needed to set the optimal price. While the Pigouvian price is optimal in the limit, it is not optimal for each realization of the VOTs with finitely many drivers.

Varying the parameters of the simulation does result in changes to the optimal level of the optimal Pigouvian price, but leaves the qualitative results of the simulation exercise unchanged: deviations from the efficient limit price are consistently associated with welfare losses. But the size of the welfare losses depends on the precise parameters of the underlying problem as well as the error in setting the efficient limit price.

5 Discussion & Conclusion

This paper shows that the traditional Pigouvian approach of internalizing social costs of congestion by setting a *single* congestion charge applies only when there are infinitely many drivers and their impact on congestion is negligible. The generally optimal solution involves charging drivers a variety of different prices. These prices depend on the desired travel time, i.e. on the number of other drivers using the same road.

One major advantage of applying mechanism design to congestion problems is that it obviates the need to conduct detailed econometric studies to estimate the distribution of the VOT. Moreover, our mechanism can be adapted to incorporate other externalities such as local- or global pollution and accidents.

A basic lesson from our mechanism is that there are potential welfare gains from introducing discrimination *within roads* with respect to travel time. We give simple examples of such pricing schedules in Section 3. Our results indicate that modern traffic management systems, such as those in Singapore or Sweden, could be improved in such a way by offering a menu of travel times with respective prices. We discuss this in more detail in the supplementary online material.

Requiring drivers to directly report their VOT may be impractical because such mechanisms might be hard to explain to people. We approach this issue by designing a mechanism in which it is a dominant strategy for each driver to reveal travel preferences truthfully. Thus, independently of what a driver believes about other traffic participants, whether correct or wrong, it is optimal to reveal her information truthfully. While this deals with potentially wrong beliefs, drivers might misunderstand a price schedule that is too complex. If this is the case, a regulator would like to design a simpler price schedule. It may be easier to use solutions that allow drivers to choose from a simplified menu of prices and arrival times. For example people could be offered a choice of three categories: fast, normal and slow. In that case they would be charged a premium for choosing faster options. The trade-offs involved in finding good user interfaces, would need to be investigated further before such congestion pricing mechanisms are implemented.

Pricing schedules could be further simplified by relaxing the implementation requirements from dominant-strategy incentive compatibility to Bayesian incentive compatibility. The price schedule of an agent would then no longer depend on the realizations of other agents but only on their distribution. This would induce continuously increasing price schedules for each driver. A caveat is that for Bayesian incentive compatibility to work, drivers need to have correct beliefs about the distribution of VOTs of other drivers.

When introducing price discrimination into a road-pricing scheme, a municipality has to deal with acceptability of a new system. In relation to a classic non-discriminating Pigouvian pricing scheme, the question who benefits and who does not depends on the distribution of VOTs. Acceptance could be encouraged in our price scheme because it allows for cross-subsidization between drivers who pay different prices. In the main model, we set the price of the outside option to be zero. But this is not necessary for incentives and a municipality could use the revenues from drivers to subsidize traffic alternatives like public transport, i.e. to set a lower price for slower means of transport. As long as price differences are unaffected, such a cross subsidy would not affect the allocation in our mechanism.

Drivers who participate in our mechanism on a regular basis might have to pay different prices each time they use these roads. Drivers might not like this uncertainty over prices. This uncertainty is already a feature of many existing traffic pricing schemes, such as the one in Stockholm or Gothenburg, but it can be eliminated by selling tickets for multiple uses or a period of time at a price equal to the expected total transfer. The mechanism would then take the VOT of that driver as constant for the time the ticket is valid. Naturally, this would lead to inefficiencies when the valuation of time varies a lot.

Before applying the results of this paper in practice one would need to extend the model to cover more complex road networks and intertemporal issues. In these cases - with a finite number of drivers - characterizing efficient allocations in simple expressions is generally impossible, since it is a discrete optimization problem. Nonetheless, efficient allocations exist in more complex networks as well, especially when the number of drivers and roads are finite. Importantly, this is not a problem of existence or implementation. Computational methods can in principle find efficient allocations and once these allocations are known, VCG-type mechanisms can be used to implement them. This should not be a major obstacle for implementation since modern traffic control systems are already today very sophisticated.

References

Arnott, R., de Palma, A., and Lindsey, R. (1990). Departure time and route choice for the morning commute. *Transportation Research Part B: Methodological*, 24(3):209–228. 1.1.1

Arnott, R. and Kraus, M. (1998). When are anonymous congestion charges consistent with marginal cost pricing? *Journal of Public Economics*, 67(1):45 – 64. 1.1.1

Arnott, R., Palma, A. d., and Lindsey, R. (1994). The welfare effects of congestion tolls with heterogeneous commuters. *Journal of Transport Economics and Policy*, 28(2):pp. 139–161. 1.1.1

Bernstein, D. and El Sanhouri, I. (1994). *Congestion Pricing with an Untolled Alternative*. Unpublished Paper. Massachusetts Institute of Technology. 1.1.1

Clarke, E. H. (1971). Multipart pricing of public goods. *Public Choice*, 11(1):17–33. 5, 1.1.2

de Palma, A. and Lindsey, R. (2011). Traffic congestion pricing methodologies and technologies. *Transportation Research Part C: Emerging Technologies*, 19(6):1377–1399. 8

Downs, A. (1993). Point of view: Implementing peak-hour road pricing at fullscale: Finding solutions to practical problems. *TR News*, (167):7–9. 1.1.1

Furuhata, M., Dessouky, M., Ordóñez, F., Brunet, M.-E., Wang, X., and Koenig, S. (2013). Ridesharing: The state-of-the-art and future directions. *Transportation Research Part B: Methodological*, 57:28–46. 6.3

Gibbard, A. (1973). Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601. 3.1

Groves, T. (1973). Incentives in teams. *Econometrica*, 41(4):617–631. 5

Jehiel, P., Moldovanu, B., and Stacchetti, E. (1996). How (not) to sell nuclear weapons. *The American Economic Review*, 86(4):814–829. 1.1.2

Jehiel, P., Moldovanu, B., and Stacchetti, E. (1999). Multidimensional mechanism design for auctions with externalities. *Journal of Economic Theory*, 85(2):258–293. 1.1.2

Kleiner, A., Nebel, B., and Ziparo, V. A. (2011). A mechanism for dynamic ride sharing based on parallel auctions. In *IJCAI*, volume 11, pages 266–272. 32

Knight, F. H. (1924). Some fallacies in the interpretation of social cost. *The Quarterly Journal of Economics*, 38(4):582–606. 1.1.1

Land Transport Authority (2016a). *Electronic Road Pricing (ERP)*. http://www.lta.gov.sg/content/ltaweb/en/roads-and-motoring/managing-traffic-and-congestion/electronic-road-pricing-erp.html (Retrieved 13 September 2016). 6.3

Land Transport Authority (2016b). *LTA to Launch Autonoumous Mobility-on-Demand Trials*. https://www.lta.gov.sg/apps/news/page.aspx?c=2&id=73057d63-d07a-4229-87af-f957c7f89a27 (Retrieved 13 September 2016). 7

Leape, J. (2006). The london congestion charge. *The Journal of Economic Perspectives*, 20(4):157–176. 6.3, 33

Mayet, J. and Hansen, M. (2000). Congestion pricing with continuously distributed values of time. *Journal of Transport Economics and Policy*, 34(3):pp. 359–369. 3, 6, 1.1.1, 2.1

Montero, J.-P. (2008). A simple auction mechanism for the optimal allocation of the commons. *The American Economic Review*, 98:496–518. 1.1.2

Morrison, S. A. (1986). A survey of road pricing. *Transportation Research Part A: General*, 20(2):87–97. 33

Mtoi, E. T. and Moses, R. (2014). Calibration and evaluation of link congestion functions: applying intrinsic sensitivity of link speed as a practical consideration to heterogeneous facility types within urban network. *Journal of Transportation Technologies*, 4(02):141. 13, 15

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73. 1.1.2, 6.2, 29

Parry, I. W., Walls, M., and Harrington, W. (2007). Automobile externalities and policies. *Journal of Economic Literature*, 45:373–399. 1, 1

Pigou, A. C. (1920). *The Economics of Welfare*. Macmillan. 1.1.1

Prud'homme, R. and Bocarejo, J. P. (2005). The london congestion charge: a tentative economic appraisal. *Transport Policy*, 12(3):279–287. 6.3

Schrank, D., Eisele, B., Lomax, T., and Bak, J. (2015). *2015 Urban Mobility Scorecard*. Texas A&M Transportation Institute and INRIX. 1

Small, K. A. (2012). Valuation of travel time. *Economics of Transportation*, 1(1):2–14. 1.1.1, 2.1

Small, K. A. and Verhoef, E. T. (2007). *The Economics of Urban Transportation*. Routledge. 18

Small, K. A., Winston, C., and Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, 73(4):1367–1382. 1, 1.1.1, 4.2, 26

Small, K. A. and Yan, J. (2001). The value of "value pricing" of roads: Second-best pricing and product differentiation. *Journal of Urban Economics*, 49(2):310–336. 1.1.1

Smith, M. and Van Vuren, T. (1993). Traffic equilibrium with responsive traffic control. *Transportation science*, 27(2):118–132. 13, 15

Steimetz, S. S. and Brownstone, D. (2005). Estimating commuters value of time with noisy data: A multiple imputation approach. *Transportation Research Part B: Methodological*, 39(10):865–889. 1, 1.1.1

The Swedish Transport Agency (2016). *Congestion Taxes in Stockholm and Gothenburg*. http://www.transportstyrelsen.se/en/road/Congestion-taxes-in-Stockholm-and-Goteborg/ (Retrieved 13 September 2016). 6.3

Transport for London (2016). *Congestion Charge*. http://www.tfl.gov.uk/modes/driving/congestion-charge (Retrieved 13 September 2016). 30

Verhoef, E., Nijkamp, P., and Rietveld, P. (1996). Second-best congestion pricing: The case of an untolled alternative. *Journal of Urban Economics*, 40(3):279–302. 1.1.1

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of Finance*, 16(1):8–37. 5

Vickrey, W. (1967). Optimization of traffic and facilities. *Journal of Transport Economics and Policy*, pages 123–136. 1.1.1

Vickrey, W. and Sharp, C. (1968). Congestion charges and welfare. *Journal of Transport Economics and Policy*, 2(1):107–125. 1.1.1

Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review*, 59(2):251–260. 1.1.1

Vickrey, W. S. (1993). Point of view: Principles and applications of congestion pricing. *TR News*, (167):4–5. 1.1.1

Yang, H., Xu, W., and Meng, Q. (2005). A sequential experimental approach for analyzing second-best road pricing with unknown demand functions. In Mahmassani, H., editor, *Transportation and Traffic Theory 2005*, pages 23–42. Elsevier. 8

6 Supplementary Material (not intended for publication)

6.1 Two Identical Roads

In this section we relax the assumption of a non-congestible outside option alternative and study the case of two identical roads. This covers the relevant case of two lanes of the same road. For example, in California there are special high-occupancy vehicle lanes parallel to normal lanes.[26] As before, there are two roads A and B. The utility of driving on each of the two roads, ignoring transfer payments, is given by $v(\theta_i, k)$. There is an odd number n of drivers, such that $n = 2m + 1$ and $m \in \mathbb{N}$. This ensures that at the efficient allocation features a fast and a slow road.[27] We refer to road A as the fast road, which means that fewer drivers use road A. The conditions on $v(\cdot, \cdot)$ are the same as before.

First-Best Allocation with Two Identical Roads We begin again by characterizing the efficient allocation that maximizes total surplus 1. We denote the efficient allocation as a function of the drivers' valuations by $x^{FB,2}(\theta)$, where the superscript 2 refers to the case of two identical roads.

Lemma 4. *Let $x^{FB,2}(\theta)$ be the efficient allocation. Then it must satisfy:*

- $\sum_i x_i^{FB,2}(\theta) < n - \sum_i x_i^{FB,2}(\theta)$.
- *For all i, j and $\theta \in \Theta$ such that $\theta_i > \theta_j$, we have $x_i^{FB,2}(\theta) \geqslant x_j^{FB,2}(\theta)$.*

The first point states that one road will be less congested than the other. This follows directly from assuming an odd number of drivers and full participation. Clearly, if all drivers are allocated to one of the roads, one road has to be faster. Since there is a slow road and a fast road, it is again optimal to put drivers with a high VOT on the fast road, starting with the highest valuations. This is the intuition behind the second point. The proof for the second point is analogous to the proof of Lemma 1. Lemma 4 implies that the faster road has the drivers with the k highest valuations and the smaller portion of drivers, i.e. $k \leqslant m$.

Next, we need to determine the efficient number of drivers on the fast road. Consider adding another driver to it. By Lemma 4 it can only be optimal to add the driver with the $(k+1)^{th}$ highest valuation to the fast road. Then welfare changes by:

$$\Delta_k^2(\theta) = \left[v(\theta^{(k+1)}, k+1) - v(\theta^{(k+1)}, n-k)\right] + \sum_{i=1}^{k} \left[v(\theta^{(i)}, k+1) - v(\theta^{(i)}, k)\right]$$
$$+ \sum_{i=k+1}^{n} \left[v(\theta^{(i)}, n-k-1) - v(\theta^{(i)}, n-k)\right]$$

The first term is the benefit for the driver that was previously using the slow road and is now transfered to the fast one. The second term reflects the loss to drivers on the fast road from sharing the road with another driver. The trade-off between these two effects is also present in the base

[26] See for example Small et al. 2005

[27] We make this assumption for simplicity. The solution for an even number of drivers is similar, although for some realizations of θ, the utility on both roads will be the same.

model without congestion on the slow road and expressed in $\Delta_k(\cdot)$. The third term is new and represents the benefit to drivers on the slow road from having one fewer driver on it. The next Lemma gives the properties of $\Delta_k^2(\theta)$.

Lemma 5. *For all $\theta \in \Theta$ and all $k \in 1, ..., m$, we have that $\Delta_k^2(\theta) > \Delta_{k+1}^2(\theta)$. Furthermore $\Delta_0^2(\theta) > 0$ and $\Delta_m^2(\theta) < 0$.*

The Proof of Lemma 5 is similar to that of Lemma 2.

Proof. We can write $\Delta_0^2(\theta) = v(\theta^{(1)}, 1) - v(\theta^{(1)}, n) + \sum_{i=2}^{n} \left[v(\theta^{(i)}, n-1) - v(\theta^{(i)}, n)\right]$. Clearly, this expression is positive. For $k = m$ we have that $\Delta_m^2 = \sum_{i=1}^{m} \left[v(\theta^{(i)}, m+1) - v(\theta^{(i)}, m)\right] + \sum_{i=m+2}^{n} \left[v(\theta^{(i)}, n-m-1) - v(\theta^{(i)}, n-m)\right]$. Given that $n - m = m + 1$ and $v_{\theta_i}(\theta_i, k+1) \leqslant v_{\theta_i}(\theta_i, k)$ this expression is negative since the first sum contains the m largest values of θ, while the second term contains the m smallest values of θ. To see that $\Delta_k^2(\theta)$ is decreasing in k, consider the first difference, which is given by:

$$
\begin{aligned}
\Delta_{k+1}^2(\theta) - \Delta_k^2(\theta) = {} & [v(\theta(k+1), k+2) - v(\theta^{(k+2)}, n-k-1)] \\
& -[v(\theta(k+1), k+1) - v(\theta^{(k+1)}, n-k)] \\
& + \sum_{i=1}^{k+1} [v(\theta^{(i)}, k+2) - v(\theta^{(i)}, k+1)] \\
& - \sum_{i=1}^{k} [v(\theta^{(i)}, k+1) - v(\theta^{(i)}, k)] \\
& + \sum_{i=k+3}^{n} [v(\theta^{(i)}, n-k-2) - v(\theta(i), n-k-1)] \\
& - \sum_{i=k+2}^{n} [v(\theta^{(i)}, n-k-1) - v(\theta(i), n-k)]
\end{aligned}
$$

Adding $v(\theta(k+2), k+1) - v(\theta^{(k+2)}, n-k)$ to line 1 and subtracting it from line 2 yields a negative term. Rearranging lines 3 and 4 as well as 5 and 6 also yields negative terms after rearranging and making use of the assumptions about $v(\cdot, \cdot)$. □

Using the results of Lemma 5 the efficient allocation is the following.

Proposition 4. *There exists some k^{**} such that for all $i = \{1, ..., n\}$ at the optimum $x_i(\theta) = 1$ if and only if $\theta_i \geqslant \theta^{(k^{**})}$. The value of k^{**} is given by:*

$$k^{**} = \max_{\Delta_k^2(\theta) \geqslant 0} k + 1 \tag{11}$$

The proof follows the same lines as the proof of Proposition 1 making use of Lemmas 4 and 5.

Implementing the First Best Having characterized the first-best allocation, we now find a payment schedule that ensures truth-telling of drivers in dominant strategies. We maximize the mechanism designer's objective function 1 subject to DIC and additionally take into account the congestion effects on the slow road. We again need to define a few new terms, similarly to Proposition

2. We denote by k^{**}_{-i} the optimal number of drivers on the fast road in the auxiliary problem when driver i's VOT is ignored, i.e. set equal to zero. We denote by $\theta^{(k)}_{-i}$ the k^{th}-highest VOT among all drivers excluding driver i.

We define the following sets:

$$\Omega_i^+(\theta) \equiv \left\{ j \neq i | \theta_j \geqslant \theta^{(k^{**})} \right\}$$
$$\Omega_i^0(\theta) \equiv \left\{ j \neq i | \theta_j \in \left(\theta_{-i}^{(k_{-i}^{**})}, \theta^{(k^{**})} \right) \right\}$$
$$\Omega_i^-(\theta) \equiv \left\{ j \neq i | \theta_j \leqslant \theta^{(k_{-i}^{**})} \right\}$$

The sets Ω_i^+ and Ω_i^0 are defined similarly as in the base model. The set Ω_i^- is the set of those drivers that are allocated to the slow road irrespective of the allocation of driver i.

Proposition 5. *With two identical roads, the first-best allocation $x^{FB,2}(\theta)$ is implemented by the following payment rule, which specifies for all $i \in 1, ..., n$:*

$$p_i^{FB,2}(\theta) = -\sum_{j \in \Omega_i^+} \left[v(\theta_j, k^{**}) - v(\theta_j, k_{-i}^{**}) \right] + \sum_{j \in \Omega_i^0} \left[v(\theta_j, n - k^{**}) - v(\theta_j, k_{-i}^{**}) \right] + \sum_{j \in \Omega_i^-} \left[v(\theta_j, k_{-i}^{**}) - v(\theta_j, -k^{**}) \right] \quad (12)$$

Proof. The proof follows along the same lines as the proof of Proposition 2. □

The first two terms are the congestion effect -$-\sum_{j \in \Omega_i^+} \left[v(\theta_j, k^{**}) - v(\theta_j, k_{-i}^{**}) \right]$ - and the reallocation effect - $\sum_{j \in \Omega_i^0} \left[v(\theta_j, n - k^{**}) - v(\theta_j, k_{-i}^{**}) \right]$ - analogous to the payments in Proposition 2. The congestion effect captures the externality on drivers on the fast road of adding driver i to this road. The reallocation effect corresponds to the externality that driver i imposes on those drivers that were on the fast road initially, but are reassigned to the slow road when driver i is considered in the optimization problem. We call the third term - $\sum_{j \in \Omega_i^-} \left[v(\theta_j, k_{-i}^{**}) - v(\theta_j, -k^{**}) \right]$- the de-congestion effect. It is new relative to equation 4. The effect represents the externality on the drivers that remain on the slow road when driver i is added to the optimization problem. The de-congestion effect has the opposite sign than the congestion effect.

When the slow road is non-congestible, moving driver i from the slow road to the fast road or adding more drivers to the slow road has no externality on drivers on the slow road. In the case of two identical roads the number of drivers allocated to the slow road now also has a congestion effect and hence affects the travel time on the slow road B. Adding more drivers to the fast road decreases the travel time on the slow road. We again normalize payments such that those drivers who do not affect the final allocation pay zero transfers. When there is congestion on both roads, this however means that drivers on the slow road might pay a positive transfer. Intuitively, drivers need to pay whenever they affect the allocation, since only in that case they impose an externality on others.

More precisely, consider a driver i who use the slow road at the efficient allocation. When driver i is not considered in the maximization, the benefit of adding another driver to the fast road is lower than when the effect on driver i is also considered. This is because when driver i is not considered, she implicitly has a value of $\theta_i = 0$, meaning she does not care about congestion. This may lead to more drivers on the slow road when i is not considered. Hence driver i might influence the final allocation with her report even though it does not affect her own road allocation. While driver i uses the slow road in both circumstances, the travel time on the slow road will be lower when she is considered. This means that drivers now pay for faster travel on *both* roads.

As before, the efficient payment schedule depends on θ_i only through its effect on k^{**} which implies that the payment schedule is again a weakly increasing step function. Also with two identical roads, the *single* Pigouvian price is not an efficient mechanism.

6.2 Revenue Maximization

Until now the mechanism designer wanted to maximize total surplus. Instead governments could use congestion charges to raise funds or a private company could be the mechanism designer. In this section the mechanism designer maximizes revenue given her beliefs about the distribution of θ. For simplicity, we assume correct beliefs $F(\cdot)$, and that $F(\cdot)$ has positive mass only on the interval $\Theta = [\underline{\theta}, \overline{\theta}]$.

We denote by $U(\hat{\theta}_i, \theta_i; \theta_{-i})$ the utility to driver i when her true type is θ_i, she reports $\hat{\theta}_i$ and the VOT of the other drivers is given by the vector θ_{-i}. Hence we have:

$$U(\hat{\theta}_i, \theta_i; \theta_{-i}) \equiv x_i(\hat{\theta}_i, \theta_{-i}) v(\theta_i, \sum_{j=1}^{n} x_j(\hat{\theta}_i, \theta_{-i})) - p_i(\hat{\theta}_i, \theta_{-i})$$

The mechanism designer thus faces the following incentive and participation constraints:

$$U_i(\theta_i; \theta_{-i}) \equiv U_i(\theta_i, \theta_i; \theta_{-i}) \geqslant U_i(\hat{\theta}_i, \theta_i; \theta_{-i}), \forall i \textit{ and } \hat{\theta}_i, \theta_i \in \Theta_i, \theta_{-i} \in \Theta_{-i}. \tag{13}$$

$$U_i(\theta_i; \theta_{-i}) \geqslant 0, \forall i \textit{ and } \theta_i \in \Theta_i, \theta_{-i} \in \Theta_{-i}. \tag{14}$$

The mechanism designer maximizes total expected revenues

$$\begin{aligned} W &= max_{[x(\cdot), p(\cdot)]} \mathbb{E}_\theta \left[\sum_{i=1}^{n} p_i \right] \\ &= max_{[x(\cdot), U(\cdot)]} \sum_{i=1}^{n} \mathbb{E}_\theta \left[x_i(\theta) v \left(\theta_i, \sum_{j=1}^{n} x_j(\theta) \right) - U_i(\theta) \right] \end{aligned} \tag{15}$$

subject to the incentive constraints 13 and participation constraints 14.

We solve this problem as in the classic optimal auction problem in Myerson (1981). The proof of the following Lemma is standard and therefore omitted.

Lemma 6. *The incentive constraints in equation 13 are equivalent to:*

$$U_i(\theta_i; \theta_{-i}) = U_i(\underline{\theta}_i; \theta_{-i}) + \int_{\underline{\theta}_i}^{\theta_i} x_i(\tilde{\theta}_i, \theta_{-i}) v_{\tilde{\theta}_i} \left(\tilde{\theta}_i, \sum_{j=1}^{n} x_j(\tilde{\theta}_i, \theta_{-i}) \right) d\tilde{\theta}_i \tag{16}$$

and

$$v_{\theta_i}\left(\theta_i, \sum_{j=1}^{n} x_j(\theta_i, \theta_{-i})\right) \tag{17}$$

is weakly increasing in θ_i.

Profit maximization requires that the participation constraint is binding for the lowest type, i.e. $U_i(\underline{\theta}_i; \theta_{-i}) = 0$ for all $\theta_{-i} \in \Theta_{-i}$. We can then insert the first expression of Lemma 6 into the mechanism designers' objective function and simplify in the usual way by using integration by parts to get the following expression:

$$max_{[x(\cdot)]} \sum_{i=1}^{n} \mathbb{E}_{\theta_{-i}} \left[\int_{\underline{\theta}}^{\overline{\theta}} \left(v(\theta_i, \sum_{j=1}^{n} x_j(\theta)) - \frac{1 - F(\theta_i)}{f(\theta_i)} v_{\theta_i}(\theta_i, \sum_{j=1}^{n} x_j(\theta)) \right) x_i(\theta) f(\theta_i) d\theta_i \right] \tag{18}$$

Substituting the *virtual value* $w_i \equiv v(\theta_i, \sum_{j=1}^{n} x_j(\theta)) - \frac{1-F(\theta_i)}{f(\theta_i)} v_{\theta_i}(\theta_i, \sum_{j=1}^{n} x_j(\theta))$, we obtain a maximization problem that corresponds to 2 subject to the additional monotonicity requirement of Lemma 6. This implies that we can use our earlier results concerning the implementation to the problem of revenue maximization. This holds as long as the allocated value of travel is weakly increasing in θ_i for all $\theta_{-i} \in \Theta_{-i}$ and as long as virtual values are declining in $k = \sum_{j=1}^{n} x_j(\theta)$, the number of drivers on the road. When $v(\theta, \cdot)$ is linear in θ_i, a sufficient condition is that virtual valuations are weakly increasing in θ_i, which is the case for example if the distribution of the VOT is such that $\frac{1-F(\theta_i)}{f(\theta_i)}$ is weakly decreasing in θ_i.

To obtain the optimal number of drivers on the road, the mechanism designer trades off the additional virtual valuation for travel time of adding another driver to the road with the reduced virtual valuation of all drivers already allocated to this road. Thus, the algorithm is the same as in Proposition 1 but using virtual valuations v_i instead of θ_i for all drivers i. We assume that valuations are distributed identically across drivers. This implies that a driver with a higher virtual value also has a larger VOT, so that a similar sorting of drivers by VOT occurs. Hence keeping the number of drivers fixed, the revenue maximizing mechanism efficiently selects drivers to drive on the road.[28] The payment schedules that implements the revenue maximizing allocation can be obtained in the usual way by substituting the resulting allocation of drivers to roads into equation 16 and rearranging.

This procedure is very similar to the one studied before, but there is an important difference: knowledge of virtual valuations requires the mechanism designer to have some belief concerning the distribution $F(\cdot)$. As a result, the revenue maximizing mechanism will only maximize revenues if the mechanism designer knows the true distribution of the VOT. Instead, the VCG mechanism is efficient irrespective of the distribution of VOTs.

6.2.1 Examples with Two Drivers Continued

We illustrate the differences between welfare and revenue maximization using the simple example with two drivers from Subsection 3.2. We now need to make an assumption about the distribution

[28] If drivers differed in their distribution of the VOT, this would no longer be true.

of the VOT of the drivers. Specifically we assume that $F(\theta_i)$ is the uniform distribution over the unit interval, so $\theta_i \sim U(0,1)$ for $i = 1, 2$. Substituting into equation 18 yields:

$$max_{[x_i(\theta)]} \sum_{i=1}^{n} \int_0^1 (2\theta_i - 1) \left[x_i(\theta)(4 - \sum_{j=1}^{n} x_j(\theta)) \right] f(\theta_i) d\theta. \tag{19}$$

A necessary condition for $x_i(\theta) = 1$ is that $\theta_i \geqslant 0.5$. If only one driver satisfies this condition, she is the only driver on the road. If both drivers satisfy the condition, the driver with the higher VOT is allocated to the road. It remains to specify when the second driver is allocated to the road. Let j be the driver with the lower VOT. The effect on profits of adding driver j to the road is given by $-(2\theta_{-j} - 1) + 2(2\theta_j - 1)$. The first term captures the virtual value of driving on the road for driver $-j$. The second term represents the additional virtual value of driver j from switching from not driving to driving. Driver j is added to the road if this term is positive. This characterizes the optimal allocation, which we depict in Figure 5:

$$(x_1^*, x_2^*)(\theta) = \begin{cases} \{0,0\} & if \quad \theta_1, \theta_2 < 0.5 \\ \{1,1\} & if \quad \theta_1, \theta_2 \geqslant 0.5 \ \& \ \theta_1 \in (0.5\theta_2 + 0.25, 2\theta_2 - 0.5) \\ \{1,0\} & if \quad \theta_1 \geqslant \max(0.5, 2\theta_2 - 0.5) \\ \{0,1\} & if \quad \theta_2 \geqslant \max(0.5, 2\theta_1 - 0.5) \end{cases}$$

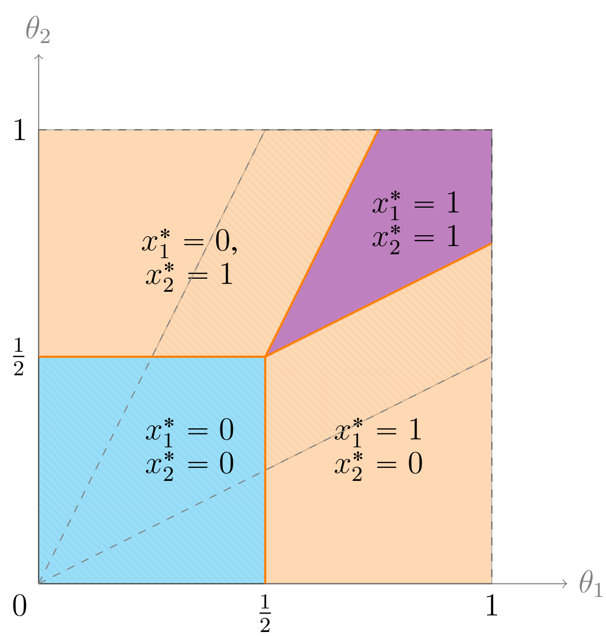

Figure 5: Revenue Maximizing Allocation with Two Drivers

As in Figure 2, the violet area shows values of (θ_1, θ_2) for which both drivers use the road. In orange areas only one driver uses the road, while in the blue area no one uses the road. The dashed lines in Figure 5 indicate the solution for the efficient allocation from Figure 2. Dashed areas highlight values of (θ_1, θ_2) in which the revenue maximizing allocation does not coincide with the efficient allocation. Notice that the number of drivers on the road is weakly lower in the revenue optimal allocation compared to the efficient allocation. Therefore the driving time on the road is also weakly lower in the revenue optimal allocation compared to the efficient allocation.

Given this allocation we can use the drivers' incentive constraints to derive the payment schedule that implements the revenue maximizing allocation:

$$p_i^*(\theta) = \begin{cases} 0 & if \quad \theta_i < \max\{0.5, 0.5\theta_{-i} + 0.25\} \\ \theta_{-i} + 0.5 & if \quad \theta_{-i} \geqslant 0.5 \ \& \ \theta_i \in [0.5\theta_{-i} + 0.25, 2\theta_{-i} - 0.5) \\ 3\max\{0.5, \theta_{-i}\} & if \quad \theta_{-i} \in [0, 0.75) \ \& \ \theta_i \geqslant \max\{0.5, 2\theta_{-i} - 0.5\} \end{cases}$$

Figure 6 shows in violet the pricing function for driver i for different values θ_{-i}. For comparison, we show in orange the payment schedule that implements the efficient allocation.

The qualitative features of the revenue maximizing payment schedule are similar to the efficient payment schedule. For values of θ_i below 0.5, the price charged by the efficient mechanism is larger than the revenue-maximizing price.[29] The mechanism designer does not allow drivers with a low VOT to use the road, because it allows her to charge higher prices for VOTs above 0.5. Note that payments are not monotone in the reported VOT of the other driver.

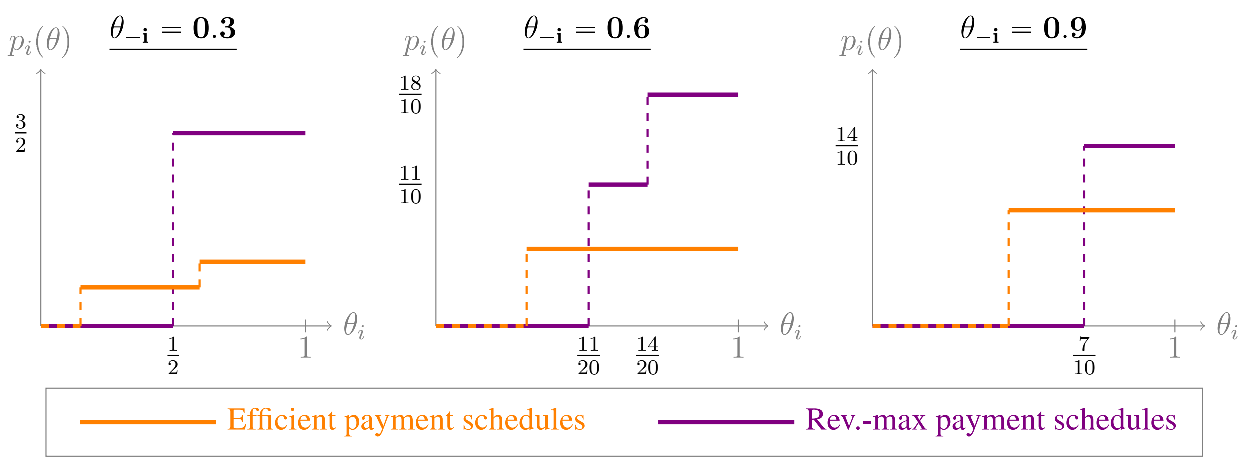

Figure 6: Payment schedules for Driver i given different values of θ_{-i}.

6.3 Congestion Pricing in Practice

In this section we discuss currently used or proposed road and congestion pricing schemes around the world and explain how they relate to our mechanism. We suggest modifications to some of these schemes that bring them closer to our proposed mechanism and thereby achieve efficiency gains. Moreover, we explore how our mechanism relates to incentives and special regulations regarding ride sharing and challenges of initial incomplete implementation.

The majority of road-pricing schemes currently in place do not contain an explicit congestion-pricing element in the sense of price discrimination with respect to the level of congestion. Many of these systems regulate long-distance traffic on highways. Take for instance per usage or per

[29]This feature also occurs in the context of auctions. When comparing the optimal auction of Myerson (1981) below the reservation price with a standard auction without a reserve price.

distance charges (e.g. road pricing for trucks in Germany, or highway fees in France and Italy, etc.) or a vignette to have the permission to use a road network (e.g. the Austrian or the Swiss highway system). These systems might indirectly affect congestion by reducing overall demand for car rides. But since prices do not adjust to congestion levels or travel time, they do not directly address congestion. Some systems deal with the congestion externality more directly by charging higher prices during rush hours or when road usage is high.

Road-pricing schemes in urban areas more frequently aim at reducing congestion. A particularly simple form of congestion pricing is *cordon area congestion pricing*, i.e. a fee to enter a specific (congestion prone) area of the city. In many cases, prices vary to achieve other objectives, for example environmental ones (e.g. in London or Milan).[30] If the fee structure of road prices is not flexible with respect to traffic conditions, the resulting level of congestion is unlikely to be efficient. Even though Leape (2006) argues that the introduction of the London congestion charge is "a triumph of economics" and led to time savings and better reliability of transport in general, he only finds small positive net benefits. Prud'homme and Bocarejo (2005) on the other hand argue that the economic benefits of the congestion charge represent less than 60% of the economic costs.

More complex and flexible systems are in place in Singapore, and the Swedish cities Gothenburg and Stockholm (The Swedish Transport Agency 2016, Land Transport Authority 2016a) and are more adequate to reducing the actual congestion externality. In these systems tolls are charged automatically, prices depend on observable characteristics and may vary over the course of the day and at different places.[31] In our mechanism, price discrimination by the VOT increases the amount and the quality of the information available. Today's existing systems are already highly complex and could easily extended to additionally incorporate unobservable characteristics, such as the VOT. In practice, the Swedish or the Singaporean systems could be extended by offering a menu of travel times with respective prices to reveal the VOT via a simple smartphone app or on-board navigation systems.

To the best of our knowledge there is no congestion-pricing system in place which uses price discrimination to gather information about the drivers' VOT. Nonetheless, the idea of our approach is partially implemented in an ad-hoc way without using monetary transfers. For instance, ambulances have a high VOT in case of accidents since any delay might cause a loss of life. In such cases, regulation gives ambulances a privileged use of roads. Similar rights are frequently awarded to convoys transporting heads of state or heavy cargo. These privileges increase travel time for other drivers. In terms of our model, these regulations are appropriate if the VOT for the privileged drivers (i.e. a convoy or an ambulance) is large enough to render the reduced travel time for other drivers efficient.

A concern regarding the practical implementation of congestion-pricing schemes is that not all drivers might participate in the mechanism right from the start. The systems in Sweden and Singapore shows that full implementation for any vehicle is feasible. But even if this was not so, we do not believe that partial implementation is a fundamental obstacle. For instance, our system could

[30]The London congestion charge for instance is a daily price (currently £11.50) for entering one zone in the inner city center which will be charged between 7:00 a.m. and 6:00 p.m.. However, there are various discounts for cars with low emissions, vans, residents, etc (Transport for London 2016).

[31]In Sweden, cameras make pictures of number plates and record place and daytime. Payments are subsequently based on these information. In Singapore, drivers are obliged to have transmitters in their vehicles and payments even vary with respect to traffic conditions, which are estimated based on retrospective data and current traffic flows, similar to a real time weather forecast.

initially be implemented only on designated fast lanes, where non-participating drivers are banned. Alternatively, non participants could be charged based on their effective travel time, assuming that they would have chosen the actual price and time combination offered by the mechanism designer. This would be similar to the system of Stockholm for non-participants. As noted earlier, charging without detailed notification is a common practice in the Stockholm (The Swedish Transport Agency 2016).

Another area of transportation research is ride sharing (see Furuhata et al. 2013 for a recent survey), which gained interest due to technological progress. Ride sharing refers to any action in which travelers share a vehicle to go from their origins to their destinations. Public transport is usually regarded as a cheap but inconvenient form due to the fixed routes and schedules. More flexible solutions could be obtained by so-called dynamic ride sharing, i.e. a real-time matching of travelers with similar itineraries. With the advent of new technologies such as GPS and smart-phones, dynamic ride sharing gained considerable commercial and academic attention.[32] Our model does not feature ride-sharing. However, the introduction of any per vehicle fees increases the incentives to share rides, which could further reduce the extent of congestion problems in practice.[33]

Many cities try to increase incentives to engage in car pooling and ride sharing, e.g. the designation of special bus lanes, which can only be used by buses and taxis. In the context of our model, this could be efficient, since a bus usually carries several people, implying a large VOT of the bus overall. Similarly, the High-Occupancy Vehicle Lanes in California allow faster travel for cars which seat more passengers. The assumption is that vehicles carrying more people implicitly have a higher VOT. In contrast to our optimal mechanism, these allocations are implemented through rigid rules rather than flexible pricing schemes. Rules such as free access to fast roads for cars with a minimum number of drivers are not necessarily desirable. Such a rule increases the number of cars on the fast road without explicitly taking valuations into account. With our mechanism implemented, such rules might no longer be needed.

[32] Kleiner et al. (2011) for instance, evaluate the potential of a VCG mechanism for the efficient matching of drivers to vehicles in such a dynamic ride-sharing setting.

[33] When Singapore introduced congestion pricing in 1975, the number of cars entering the center decreased by 41.6% initially and 22.9% in the long-run, (Morrison 1986). Interestingly, Leape (2006) does not report any effects of the London congestion charge on car pooling.