Internalizing the hypercongestion externality on highways

Kathrin Goldmann*
Gernot Sieg*
May 10, 2017

Abstract

Traffic jams on highways can occur even without bottlenecks, simply because of interaction between vehicle drivers on the road. From an aggregated point of view, the instability of free flow arises stochastically, and the probability of a traffic jam increases with the saturation of the highway. We offer a method to internalize this “hypercongestion” externality through a Pigouvian congestion charge. Applying the method to selected German highway sections results in a maximum charge of up to 17 Euro-Cents per km, including about a 11 Euro-Cent per km hypercongestion adjustment.

JEL: L91, R41
Keywords: hypercongestion, road pricing, stochastic capacity, congestion charge

*Westfälische Wilhelms-Universität Münster, Institut für Verkehrswissenschaft, Am Stadtgraben 9, 48143 Münster, E-Mail: kathrin.goldmann@wiwi.uni-muenster.de

1
1 Introduction

The idea of congestion pricing is to implement a road user charge to make travellers fully aware of the congestion externality they impose on others (Vickrey, 1963; Walters, 1961). Speed-flow functions are backward-bending (see Figure 1); for some flows there is a decreasing branch (negative externality) that depicts congestion, but also a branch which is called hypercongestion (Small and Chu, 2003; Button, 2004) where speed is positively related to traffic flow (positive externality). Verhoef (1999) shows that hypercongestion is dynamically infeasible when considering capacity as deterministic. Deterministic capacity of a freeway is defined as the maximum flow rate that can reasonably be expected to traverse a facility under prevailing roadway, traffic and control conditions (HCM, 2000). Traffic jams can be caused by bottlenecks triggered by traffic accidents, construction sites, on-ramps, tunnels or inhomogeneous road design. However, as Sugiyama et al. (2008) showed, traffic jams even occur without bottlenecks. It is sufficient that drivers on a highway interact with each other to make the free flow unstable. There may be deterministic reasons like tailgating, to fast reaction to speed changes, slow overtaking by trucks, slow reaction because of inattentiveness or queue-jumping, but in the system, these driving errors occur stochastically (Schönhof and Helbing, 2007). Some of these factors culminate in a traffic jam, but some do not. The probability of their causing a traffic jam increases with the saturation of the highway. Therefore, Elefteriadou et al. (1995) and Brilon et al. (2005) understand capacity as stochastic.

A stochastic capacity approach enables us to integrate hypercongestion into the static economic model of road congestion pricing. For some flow rates, there are two types of speed, congested and hypercongested, and the probability as to which of the speeds prevail depends on the flow rate. A driver entering the road to travel a certain distance faces a stochastic travel time, depending on the number of other cars on the road.

The driver only considers his own costs, but not the time losses other drivers have due to increased traffic. The Pigouvian congestion charge internalizes this congestion externality. We identify a second externality of an additional driver
on the road; the driver increases the probability of hypercongestion. In a state of hypercongestion, average speed is low and therefore time losses are large. We show that the second effect, which has so far been ignored when calculating Pigouvian congestion charges, is about as important as the usual congestion effect of travel time prolongation, by calculating the Pigouvian hypercongestion charge for segments of the German Autobahn.

2 The hypercongestion externality

Empirical speed flow charts (see Figure 1) usually exhibit two branches, an upper one on which speed decreases with flow due to congestion, and a lower one on which speed increases with flow. The lower branch is called hypercongestion in economics, whereas engineers, traffic planners or traffic police departments only consider hypercongestion as traffic jam. Hills (1993) underlines the importance of considering the strong bimodality in the distribution of the average speeds for constructing a congestion cost function, but does not propose a method for implementing it.

Verhoef (1999) shows that hypercongestion is dynamically infeasible when considering capacity as deterministic. In order to depict hypercongestion in a static model with continuous demand, inflows onto the road should exceed the maximum possible inflow at some point in the past, which is inconsistent with the concept of maximum deterministic capacity. Small and Chu (2003) suggest that hypercongestion on a highway entails a queue of cars waiting in front of a bottleneck. Therefore, density within the queue does not have an effect on the outflow rate from the bottleneck and on travel time, but only on the number of cars waiting. Following this interpretation, Small and Chu (2003, p. 326) state that hypercongestion is irrelevant to users who care only about total travel time.

A few papers have used the bottleneck model (Vickrey, 1969; Arnott et al., 1990; Small, 2015) for analyzing hypercongestion, by postulating that bottleneck capacity varies with the length of the queue. Yang and Huang (1997) identify a dynamic externality, i.e. how an additional car influences the queue length and
Figure 1: Speed $v$ - flow $q$ relationship

Source: Own representation based on speed-flow data of the highway A45/1GIN from Brilon and Geistefeldt (2010).
therefore the bottleneck efficiency, but do not consider inefficient hypercongestion states such as stop and go traffic, but only queues. Consequently, Yang and Huang (1998) include a queuing externality in the congestion charge and suggest calculating the flow-dependent travel time and the queuing delay separately, with the former being predicted by an analytical delay formula and the latter determined from network equilibrium conditions.

Elefteriadou et al. (1995) analyze empirically ramp-freeway junctions and establish that breakdown is a probabilistic variable and showed that breakdowns do not always occur at specific volumes, even at the same site. Spontaneous formations of jams have also been observed in the absence of any bottleneck. For example, Treiterer and Myers (1974) have depicted the trajectories of a phantom jam. Coifman (1997) also presented the trajectories of thirteen shockwaves in the morning rush hour, where a small disturbance grows as it propagates upstream until vehicles come to a standstill. Sugiyama et al. (2008), Nakayama et al. (2009) and Tadaki et al. (2013) performed traffic experiments on a circuit to investigate the emergence of a jam without a bottleneck. Their experiments showed that when the density is below a critical value, traffic flow is unconditionally stable and no jam emerges (e.g. $q < 2,400$ in Figure 1).\footnote{Due to the fundamental relationship between mean speed, density and flow, densities can be converted into flows.} Above this value, a random jam formation can be observed. Whereas the mechanism of traffic flow breakdown remains controversial in the literature, a consensus is reached regarding the stochastic nature of traffic breakdown occurrences (Wang et al., 2010).

The bathtub model (Arnott, 2013; Fosgerau and Small, 2013; Fosgerau, 2015; Arnott et al., 2016) analyzes urban hypercongestion. A backward-bending fundamental diagram of traffic flow also applies at the level of an urban neighborhood, which meets certain conditions (Daganzo, 2007; Geroliminis and Daganzo, 2008; Daganzo et al., 2011). As a result, urban congestion can be analyzed in an aggregate manner, using a speed-flow relationship as we do for highway-congestion. In contrast to cities, where trips begin and end at various locations, on highways, routes are unidirectional and cars are somehow caught in a tube. Accordingly, a
driver who makes a trip of a certain length in a city influences the speed of other cars, independent of where the trip begins or ends. However, whereas in a city even cars that start later, influence cars already on the road, in our model for highways, traffic is not influenced by cars starting later. While Fosgerau (2015) identifies large hypercongestion costs in urban traffic, we identify hypercongestion costs for non-urban unidirectional traffic flows.

Hypercongestion can surely occur “as a transient response of a non-linear system to a demand spike” (Arnott, 1990, p. 200), or as a transient reduction in capacity (due to a traffic accident), or as a queue in front of a bottleneck (Small and Chu, 2003). All these approaches consider capacity as deterministic. We focus on hypercongestion that occurs in a non-linear system without identifiable reasons and therefore assume, as do Elefteriadou et al. (1995), Brilon et al. (2005) and Wang et al. (2010), that capacity is stochastic.

The model allows for calculating the costs of congestion and hypercongestion for single highway sections. We are able to compare the extend of time losses and travel time costs split for congestion and hypercongestion over the course of a day and identify different patterns for the two externalities. The method, however, simplifies some aspects.

We directly use the fundamental diagram and traffic flow data that displays the number of vehicles per hour passing a specific traffic detector. We do not explicitly make a statement about how drivers enter the highway. Highway sections can either be without on-ramps or with on-ramps that provide small traffic inflows that do not affect the stability of traffic flow. Sections, where on-ramps provide large traffic inflows, cannot be analyzed with this model. We focus on highway sections where hypercongestion occurs due to driving behavior and not due to transient demand or supply spikes as well as bottlenecks. Moreover, the model is a static model that allows for calculating congestion and hypercongestion charges at a specific point of the highway. Traffic flow dynamics cannot be displayed. When calculating the congestion charges we assume that a specific traffic situation and accordingly the calculated charges hold for one kilometer around the traffic detector.
 Whereas these simplifications limit the theoretical foundation of the model, they have advantages for practical implementation. When information about the speed-flow relationship and the breakdown probability is available, we demonstrate, that the charge can easily be calculated with average hourly traffic flow data.

3 Internalizing congestion and hypercongestion

The expected travel speed of a driver depends on the traffic situation on the highway. Traffic can either be congested at an average travel speed of $\bar{v}(q)$ or hypercongested at a speed of $v(q)$. Figure 2 depicts a typical speed-flow relationship for motorway sections where hypercongestion prevails.

The probability of an unstable hypercongested traffic situation is $p$, so that the
expected travel speed can be written as

\[ E(v) = pv + (1 - p)\bar{v}. \]  

The speed \( v \) and the breakdown-probability \( p \) in this case both depend on flow \( q \), measured by the number of vehicles per hour. In a situation of bottlenecks, the converse may be true, meaning that speed influences flow (Daganzo, 1997). The marginal effects that an additional driver imposes on subsequent drivers can be written as

\[
\frac{dv}{dq} = (1 - p)\frac{d\bar{v}}{dq} - \frac{dp}{dq} \left( \bar{v} - v \right) + p\frac{dv}{dq} = \frac{d\bar{v}}{dq} - \frac{dp}{dq} \left( \bar{v} - v \right) - p \left( \frac{d\bar{v}}{dq} - \frac{dv}{dq} \right). 
\]  

(2)

Given the probability \( 1 - p \) that the traffic flow remains stable, we have the normal congestion effect. Additional drivers who enter the road decrease the speed of the whole system. The second effect describes the situation where traffic has already broken down and an increase in cars per hour passing that point can be interpreted as a marginal recovery from hypercongestion. The third effect is the drop from stable to hypercongested traffic. This implies a speed loss of \( \bar{v} - \bar{v} \).

Because the expected private travel time costs \( C \) of a driver depend on his average speed \( v \), we have

\[ C(q) = p \cdot c(v) + (1 - p)c(\bar{v}) \]  

and because all drivers are identical, these costs are the average costs of each driver \( q \) on the road. Social costs are \( SC = q \cdot C(q) \) and marginal social costs are \( MSC = C + qC' \). A driver internalizes \( C \), but \( qC' \) is the external effect a Pigouvian tax is determined to internalize. The marginal social travel time costs are:

\[
q\frac{dC}{dq} = p \cdot \frac{dc}{dv} \frac{dv}{dq} + \frac{dp}{dq} c(v) + (1 - p) \cdot \frac{dc}{dv} \frac{\bar{v}}{dq} - \frac{dp}{dq} c(\bar{v}). 
\]  

(4)

Considering a distance of \( a \) and a time value of \( t \), \( c(v) = ta/v \) and \( dc/dv = -ta/v^2 \), equation 4 can be written as:

\[
q\frac{dC}{dq} = -p \frac{ta}{v^2} \frac{dv}{dq} + \frac{dp}{dq} \frac{ta}{v} + (1 - p) \cdot \frac{(-ta)}{\bar{v}^2} \frac{d\bar{v}}{dq} - \frac{dp}{dq} \frac{ta}{\bar{v}}. 
\]  

(5)
To summarize, the hypercongestion charge, or the congestion charge 2.0, equals
\[
\frac{dC}{dq} = -qta \left[ \frac{1}{v^2} \frac{dv}{dq} \right] -qta \left[ p \left( \frac{1}{v^2} \frac{dv}{dq} - \frac{1}{\bar{v}^2} \frac{d\bar{v}}{dq} \right) + \frac{dp}{dq} \left( \frac{1}{\bar{v}^2} - \frac{1}{v^2} \right) \right]. \tag{6}
\]

The first term equals the traditional Pigouvian congestion charge (or Congestion Charge 1.0), but to address hypercongestion, an adjustment described by the second term is needed. To distinguish between the three effects shown in Figure 2, equation 6 can be rearranged to
\[
\frac{dC}{dq} = -qta \left[ \left( 1 - p \right) \left( \frac{1}{v^2} \frac{dv}{dq} \right) + p \left( \frac{1}{\bar{v}^2} \frac{d\bar{v}}{dq} \right) + \frac{dp}{dq} \left( \frac{1}{\bar{v}^2} - \frac{1}{v^2} \right) \right]. \tag{7}
\]

As the derivative \( dv/dq \) is negative, multiplication by the total costs \((-qta)\) increases marginal social costs. On the lower branch, additional drivers increase the flow \((dv/dq > 0)\) and for this reason, marginal social costs decrease. For the third effect, the derivative \( dp/dq \) is positive and the term in brackets will always be negative, as the travel time on the lower branch is greater than on the upper branch. For this reason, this effect also increases marginal social costs.

### 4 Application for selected highways

To calculate the marginal speed losses (upper branch), speed gains (lower branch) as well as speed drops (distance between upper and lower branch) for the congestion charge, we need the functional forms for the speed-flow relationships for selected German motorways.

The fundamental relationship describes the relation between flow \( q \), density \( k \) and space mean speed \( v \). Taylor et al. (2008) compare the performance of eight different functional forms in modelling different traffic regimes. They find that each model has certain advantages in representing specific traffic regimes, but fails to represent others. Rakha and Crowther (2002) compare Greenshield’s single-
regime, Pipe’s two-regime and Van Aerde’s single-regime model. They demonstrate the shortcomings of Greenshield’s and Pipe’s models in capturing the entire range of traffic stream situations. They find that the four-parameter Van Aerde model is able to reflect different traffic situations on different road types, as it best approximates the field data. Van Aerde’s (1995) model describes the speed-density relationship by means of the minimum distance headway between consecutive vehicles. In a stable relationship between traffic density, traffic flow and speed, the basic relationship can be written as

\[ q(v) = \frac{v}{c_1 + c_2/(v_0 - v) + c_3v}. \]  

(8)

where \( c_i \) are parameters of the function and \( v_0 \) is the speed at a flow or density of zero. The van Aerde (1995)-function is backward-bending and each value of \( q \) can be assigned to one speed of congested \( v \), as well as to one speed of hypercongested \( v \) traffic. Brilon and Geistefeldt (2010) analyze traffic flows on German highway segments in order to revise the design capacities. They found that the Van Aerde model provides the best fit for highway sections where hypercongestion occurs and publish the four parameters needed to specify the function. For this reason, we use the results of their traffic flow analysis to calculate the congestion charge.

Coming from very low traffic flows corresponding to high speeds, the more cars on this highway section per hour, the more the probability increases that the traffic will break down and speed will drop to the lower part of the van Aerde function in the speed-flow diagram in Figure 2. It is widely accepted in the literature that the breakdown flow/density has properties of a random variable (Elefteriadou et al., 1995; Lorenz and Elefteriadou, 2000; Brilon et al., 2005). Wang et al. (2010) define 12 traffic states by means of density intervals of constant size and calculate a transition probability matrix that a specific future state will follow the present state for a highway in the US state of Georgia. They find that the breakdown probability in this discrete model follows a Zipf distribution. The transition probability matrix also allows for an analysis of recovery from hypercongested traffic states. Chow et al. (2009) develop a bivariate Weibull distribution for which the breakdown probability is a function of both mean speed and the occupancy of approaching traffic. They calculate its parameters for a freeway section in California.
and suggest that the results can be advantageous in traffic management and control applications. Focusing on highway capacity analysis, Brilon et al. (2005) and Brilon and Geistefeldt (2010) found that the normal Weibull distribution best fits the breakdown properties of the investigated German motorway sections. For this reason, the latter approach has been adopted in this paper for congestion charge calculation. The distribution function of the Weibull distribution has the following form:

\[
F(q) = 1 - e^{-\left(\frac{q}{\beta}\right)^\alpha}.
\]

(9)

For each additional driver using the road per hour, the probability of a breakdown increases marginally. As breakdowns of traffic flows occur suddenly, only short time intervals are appropriate for analysing traffic break-downs. Brilon and Geistefeldt (2010) use five-minute intervals to estimate the parameters of the Weibull distribution and convert them back to hourly intervals by assuming that the variance of the traffic flow is normally distributed over the interval. Knowing the functional forms of the speed-flow relationships and the breakdown probabilities, we are able to calculate the costs of the three identified externalities.²

For illustration, the Pigouvian congestion charge has been calculated for four German motorway sections (A3 northbound, A42 eastbound, A42 westbound, A45 northbound) which are located in the states of North Rhine-Westphalia and Hesse (see Figure 3). The sections differ regarding the number of lanes, their location (inside or outside metropolitan areas) and the presence of speed limits (see Table 1). The respective traffic detectors are roughly indicated on the map, while the precise GPS position is given in Table 1.

The congestion and hypercongestion charge both depend on the traffic flow \(q\). The traffic flow per hour and weekday in 2015 is taken from The Federal Highway Research Institute (BASt, 2017). We calculate the annual average flow per weekday and hour, where weekdays exclude school or public holidays.

²The input values for the Van Aerde function and the Weibull distribution are presented in Table A.1 in the appendix. The parameters are from Brilon and Geistefeldt (2010) and were calculated with traffic data from 2001 to 2006.
It should be noted that the size of the congestion charge depends on the travel time costs used for calculation. We differentiate between three different travel time cost categories. There are private trips (for shopping, leisure activities or driving to the workplace and back), business trips and trips of heavy-duty vehicles. The German methodology handbook for the federal infrastructure plan differentiates between private and business time costs parameters, while both increase with the total trip length (BMVI, 2016). The study “Mobility in Germany” contains the average car trip lengths, as well as the trips broken down by purpose (Infas and DLR, 2008). However, the results include all trips and not just those on highways. For this reason, we made the assumption that the average private trip length of
Table 1: Characteristics of selected highway sections

<table>
<thead>
<tr>
<th>Highway No.</th>
<th>A45 northbound</th>
<th>A3 northbound</th>
<th>A42 eastbound</th>
<th>A42 westbound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>rural area</td>
<td>metropolitan area</td>
<td>metropolitan area</td>
<td>metropolitan area</td>
</tr>
<tr>
<td>Highway km</td>
<td>179.1</td>
<td>166.5</td>
<td>45.7</td>
<td>45.7</td>
</tr>
<tr>
<td>Number of lanes</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Speed limit</td>
<td>none</td>
<td>none</td>
<td>100 km/h</td>
<td>100 km/h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Position of traffic detectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>WGS84E</td>
</tr>
<tr>
<td>WGS84N</td>
</tr>
</tbody>
</table>

approximately 15 km is somewhat higher for trips on highways (45 km). This assumption is necessary, as the time value function is upward sloping with the trip length. The corresponding time costs are 8.17 Euro/h. The average trip length of business trips on highways is assumed to be 100 km (time costs: 30 Euro/h).

There are basically two types of heavy-duty vehicles on the road: normal trucks and semi-trailer trucks. Due to different trip lengths and vehicle specifications, the drivers wages (17.64 and 20.14 Euro/h) and the capacity maintenance costs (5.81 and 9.34 Euro/h) differ. Moreover, the methodology handbook also offers an average time value for transported goods of 6.88 Euro/h with an average loading factor of 0.7 (BMVI, 2016). The total time costs for normal trucks are therefore assumed to be 28.27 Euro/h and 34.30 Euro/h for semi-trailer trucks. On selected highways, among heavy duty vehicles, the shares of normal trucks versus semi-trailer trucks are approximately 2/3 versus 1/3, which yields an average time cost value for heavy duty vehicles of 30.28 Euro/h.\(^3\)

The shares of trips by purpose are also from the Mobility in Germany study, although the purposes of trips including routes on highways may differ from trips

\(^3\)Another cost component, which was not included due to complexity, is fuel consumption. A shift from flowing to stop+go traffic (3. effect in Figure 2) will increase fuel consumption and is therefore a negative externality imposed on the other drivers as well.
within urban centers. However, detailed data for highway trips is not available. The same applies to the average rate of vehicle occupancy $r_{VO}$ of 1.1 (Infas and DLR, 2008). The cost factors are weighted by the share of private ($w_p$), business ($w_b$) and heavy duty vehicle ($w_{hd}$) trips.

It should be noted that the first identified congestion effect in Figure 2 does not affect heavy-duty vehicles, as their allowed maximum speed is 80 kilometers per hour and therefore, they do not incur travel time prolongation on the upper branch. The travel time cost parameter $c_1$ is:

$$c_1 = w_p \cdot r_{VO} \cdot 8.17 \, \text{€} + w_b \cdot r_{VO} \cdot 30 \, \text{€} = 11.80 \, \text{€},$$

where the weights are $w_p = 0.88$ and $w_b = 0.12$.

The second and third congestion effect is relevant for all vehicles on the highway, including heavy-duty vehicles. For this reason, the weights are somewhat different at $w_p = 0.77$, $w_b = 0.10$ and $w_{hd} = 0.125$.

$$c_{2-3} = w_p \cdot r_{VO} \cdot 8.17 \, \text{€} + w_b \cdot r_{VO} \cdot 30 \, \text{€} + w_{hd} \cdot 30.28 \, \text{€} = 14.12 \, \text{€}$$

Evaluating the travel time losses due to normal congestion (1. effect) and the travel time losses due to hypercongestion (2. and 3. effect) with the cost parameters, enables us to calculate a congestion charge that depends on the current traffic flow situation.

Figure 4 shows the congestion charge due to marginal speed losses on the upper branch (blue area), as well as the hypercongestion adjustment (red area) for an average Thursday in 2015. It is evident that in peak times, due to the increase in probability of a traffic breakdown, the costs of a shift from the upper to the lower branch become more pronounced. More precisely, when the flow exceeds 60% of design capacity flow, the hypercongestion charge starts to rise. Due to the higher costs of a traffic breakdown, compared to the marginal speed losses on the upper branch, the hypercongestion charge increases faster that the normal congestion charge. In off-peak periods, the probability of a random breakdown is close to zero, as driver errors that do not result in bottlenecks such as accidents, do not affect the stability of the traffic flow up to a certain saturation level and therefore,
(a) A3 northbound
(b) A45 northbound
(c) A42 westbound
(d) A42 eastbound

Figure 4: Congestion charge 1.0 (blue) and Hypercongestion adjustment (red)
the hypercongestion externality is zero as well. Traffic breakdowns at low traffic flows are caused by bottlenecks and as mentioned in Section 1, traffic breakdowns because of bottlenecks have to be analyzed with different models.

The maximum charge for the A3 (see Figure 4a) is about 17 Euro-cents/km, including about 11 Euro-cents/km hypercongestion adjustment. Ignoring hypercongestion by using only the upper branch of the speed-flow-relationship to calculate the congestion charge leaves aside two thirds of congestion effect in the case of the A3. The average charge per km No. 167 on the A3 on an average Thursday in 2015 is about 3.7 Euro-Cents, whereas the average charge per km throughout the year, including weekends and public holidays, is 1.7 Euro-Cents per km.

When drivers are confronted with traffic delays, they can change their route (reassignment), their destination (redistribution), their time of departure (rescheduling), their mode (mode-switching), their vehicle-occupancy (car-sharing/pooling) or the frequency of their trip-making, or opt out entirely (trip suppression) (Hills, 1993). Demand is elastic and the Pigouvian charge internalizes the external effect at the equilibrium, where demand is lower than without the charge. Because no data is available to determine the demand function for the highway sections we analyze, the values we calculate are only accurate if demand is inelastic, i.e. if drivers do not change their behavior. However, demand should be elastic when implementing a congestion charge. If there is no reassignment to the peak period of the considered highway, then the peak charge of about 17 Euro-cents per kilometer is the upper bound of the congestion charge. If there is rescheduling, then congestion and the congestion charge can rise in previously non-peak periods.

However, when considering congestion charges, some caveats are always necessary (Evans, 1992). A congestion charge may redistribute welfare to the rest of society, if the revenue is not earmarked for highways. Furthermore, the size of the welfare gain may be much smaller than the redistributed charges, such that road users will lose massively. If peak times are becoming costly, users who so far paid with longer journey times, then have to pay with money. If richer drivers can shift their travel times more easily than poorer drivers, the charge will be indirectly regressive. And finally, congestion charge revenues are higher if capacity is
sub-optimally small. For a revenue-maximizing government therefore, congestion charges are a perverse incentive for supplying road capacity.

5 Conclusion

Due to financial constraints confronting public authorities, as well as decreasing implementation and transaction costs for distance-based road charges, their implementation may increase in future. Therefore, the accurate calculation of road charges is of great importance. Charges can be calculated in different ways (Newbery, 1988). If politicians opt to implement a congestion charge on highways, each highway section has to be analyzed regarding its characteristics. If it is a bottleneck, or a queuing area in front of a bottleneck, it is reasonable to use the approach of Yang and Huang (1998) to determine the charge. However, if the part of the highway (also) suffers from stochastic traffic breakdowns without bottlenecks, we recommend using the Pigouvian Hypercongestion Charge developed in this paper.

We distinguish between a congestion and a hypercongestion effect. A traditional Pigouvian congestion charge only internalizes the incremental time losses of drivers on the road, but ignores the fact that additional drivers also increase the probability of stop-and-go traffic. Therefore, the traditional Pigouvian charge is too low. We show that when traffic flow exceeds about 60% of design capacity flow, the hypercongestion charge becomes relevant on selected German highway sections. At rush hour, the traditional congestion charge only internalizes approximately half of the external travel time costs imposed on other drivers.

The maximum peak charges calculated are only an upper bound for the equilibrium charges after road users have adapted to the new system. Because demand is elastic, the equilibrium charges are smaller at the peak period, but the charged peak period may expand.

Our model is a static model based on average speed-flow data which is easily available, at least in Germany. If the speed-flow function and the breakdown probabilities are known, the calculation of the charges is comparatively simple. One drawback of this model is, however, that we cannot analyze transitions between
different traffic regimes like e.g. performed by Daganzo (2011).

Another obstacle to the implementation is that because trucks are longer, slower, and heavier, they warrant a special treatment and should receive a special charge (Verhoef et al., 1999). Based on Coifman (2015) who identifies that many of the critical parameters of the flow-density relationship depend on vehicle length further research should separate the external effects of trucks and cars on travel times to determine vehicle type dependent charges that internalizes congestion and hypercongestion.
References


19


### Table A.1: Parameters and values used for application

<table>
<thead>
<tr>
<th>Parameter</th>
<th>A45 northbound</th>
<th>A3 northbound</th>
<th>A42 eastbound</th>
<th>A42 westbound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>0.004847</td>
<td>0.004303</td>
<td>0.008312</td>
<td>0.007586</td>
</tr>
<tr>
<td>$c_2$</td>
<td>0.38611</td>
<td>0.11140</td>
<td>0.27931</td>
<td>0.24128</td>
</tr>
<tr>
<td>$c_3$</td>
<td>0.0001276</td>
<td>0.0000753</td>
<td>0.0000469</td>
<td>0.0000739</td>
</tr>
<tr>
<td>$v_0$</td>
<td>138.9</td>
<td>130.14</td>
<td>117.1</td>
<td>116.7</td>
</tr>
<tr>
<td>$q_{\text{max}}$</td>
<td>3721</td>
<td>6500</td>
<td>4084</td>
<td>3986</td>
</tr>
</tbody>
</table>

| Weibull-distribution$^a$          |                |              |               |               |
| $\alpha$                         | 12.55          | 16.62        | 10.93         | 15.37         |
| $\beta$                          | 4225           | 6607         | 4571          | 4252          |

| Travel time cost parameters in Euro$^b$ | | | | |
| upper branch (1. effect)            | 11.80          |               |               |               |
| lower branch (2. effect)            | 14.12          |               |               |               |
| break down (3. effect)              | 14.12          |               |               |               |

$^a$ From Brilon and Geistefeldt (2010).

$^b$ Own calculations based on Infas and DLR (2008) and BMVI (2016).