Westfälische
Wilhelms-Universität
Münster

Quantifying the phantom jam externality: The case of an Autobahn section in Germany

by
Kathrin Goldmann and Gernot Sieg

Institute of Transport Economics Münster
Working Paper
No. 30
March 2020

wissen.leben
WWU Münster

Address
Institut für Verkehrswissenschaft
Am Stadtgraben 9
D 48143 Münster, Germany

Telephone
+49 251 83-22 99 4

Fax
+49 251 83-28 39 5

E-Mail
verkehrswissenschaft@uni-muenster.de

Website
http://www.iv-muenster.de

The Working Paper Series seeks to disseminate economic research work by the WWU, Institute of Transport Economics staff and visitors. Papers by researchers not affiliated with the WWU Institute of Transport Economics may also be considered for publication to the extent that they have been presented at research seminars/workshops organised by the institute.

The working papers published in the Series constitute work in progress circulated to stimulate discussion and critical comments. Views expressed represent exclusively the authors' own opinions.

The Series is managed by the Director of the Institute of Transport Economics.

Quantifying the phantom jam externality: The Case of an Autobahn section in Germany

By Kathrin Goldmann* and Gernot Sieg*

When traffic demand is high, traffic jams can occur in the absence of bottlenecks or demand peaks, simply because of the interaction of vehicle drivers on the road; a phenomenon called phantom jam. The probability of phantom jams occurring increases with traffic flow. Road users only consider their own time costs and not those of other drivers, so that an unpriced phantom jam externality leads to inefficient road usage. We offer a method for quantifying the phantom jam externality and apply the method to a specific highway section in Germany. Congestion charges calculated ignoring phantom jam externalities may be up to 50 percent too low.
JEL: L91, R41
Keywords: hypercongestion, congestion costs, stochastic capacity, phantom jams, external costs

I. Introduction

Traffic congestion during the rush hour remains an observable phenomenon worldwide. It results in significant travel time losses for commuters, additional external environmental costs and a loss of attractiveness of the affected areas. Reasons for congestion on highways can be on the demand side (on-ramps with high inflows or fluctuations in demand) and on the supply side (traffic accidents, construction sites, tunnels, inhomogeneous road design or simply insufficient capacity).

Besides these deterministic reasons, Sugiyama et al. (2008), Nakayama et al. (2009) and Tadaki et al. (2013) show that traffic jams can also occur randomly due to driving behavior. When traffic density exceeds a critical value, phantom jams may occur even in the absence of supply side reasons. Although density remains constant in their experiment, traffic is freely flowing initially, but breaks down after a while. To make the initial free flow unstable, it is sufficient that drivers on a highway merely interact with each other. For each phantom jam, there may be a deterministic reason like tailgating, excessively fast driver reactions to speed changes, slow overtaking by a truck, slow reactions because of drivers inattentiveness or queue-jumping, but in the system, these driving errors occur stochastically and may or may not culminate in a traffic jam (Schönhof and Helbing, 2007). The probability of their causing a traffic jam increases with the saturation of the highway. For this reason, capacity cannot be considered as a fixed value, but seen rather as a stochastic concept (Elefteriadou et al., 1995; Brilon et al., 2005).

Economic congestion models[1] can be classified as bottleneck, bathtub and

* Westfälische Wilhelms-Universität Münster, Institute of Transport Economics, Am Stadtgraben 9, 48143 Münster, Germany.

[1] Whereas economists refer to the traffic state represented by the upper branch in a speed-flow diagram as congested, because this traffic state already imposes marginal speed losses on other drivers (externalities), other certain sciences consider this traffic state as freely flowing traffic. For economists, only the small horizontal part of the upper branch is free flow, because there is no externality. Given that, for our analysis, the section without externalities is negligibly small, we use

speed-flow models. The bottleneck model (Vickrey, 1969; Arnott et al., 1990, 1993; Small, 2015; Xiao et al., 2015) allows for analyzing hypercongestion (jammed traffic), by considering queuing delays in front of a bottleneck with fixed or stochastic capacity. Different tolling systems can be evaluated regarding their ability to eliminate queuing in front of the bottleneck. The bathtub model (Arnott, 2013; Fosgerau and Small, 2013; Fosgerau, 2015; Arnott et al., 2016) analyzes urban hypercongestion at an aggregate level. In the morning rush hour, cars enter the downtown urban center and when density is sufficiently large, traffic flow becomes inefficiently low and the outflow of cars decreases, which makes hypercongestion more persistent. A time-varying toll or traffic management systems should therefore avoid hypercongestion. Some of the above mentioned models involve simulations. Generally, simulations are also performed by engineers for specific roads and road networks. For instance, He et al. (2016) optimize the simulation framework for an optimal time-varying pricing of toll roads. The results enable, for example, the evaluation of toll adjustments regarding their impact on changes in demand, length of peak periods or toll revenue.

Speed-flow models directly use the fundamental diagram to analyze congestion and hypercongestion. However, Verhoef (1999) shows that in speed-flow models, 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 must have exceeded the maximum possible inflow at some point in the past, which is inconsistent with the concept of maximum deterministic capacity. Moreover, Verhoef (1999, 2005) shows that for roads without a downstream bottleneck, the average cost curve is backward-bending, and intersections with the demand curve yield multiple and unstable equilibria.[2]

In contrast, traffic engineers still use speed-flow models to determine the capacity of highways, for instance in Highway Capacity Manuals. To incorporate the fact that road capacity is not a fixed value, Brilon et al. (2005) and Brilon et al. (2007) and Brilon and Geistefeldt (2010) show how to implement the stochastic nature of traffic flow breakdowns.

We revisit speed-flow models by incorporating stochastic road capacity. We calculate the expected costs of congested and hypercongested traffic states. In our application, we obtain a dynamically consistent average cost curve that is not backward-bending. This is due to the fact that the probability of costly traffic hypercongestion occuring, increases with flow as well.

As we only consider congestion costs under prevailing traffic-flow conditions, we are not able to analyze demand reactions, because traffic-flow cannot be equated directly to demand. For this reason, we do not offer a complete economic model that allows for analyzing welfare gains due to congestion charges. However, speed-flow data offers a very precise description of traffic situations and is available for various road sections in developed countries. We do not need to make assumptions

the two terms synonymously and refer to the upper branch as congested or as freely flowing traffic. Economists refer to the traffic state represented by the lower branch as hypercongestion, whereas other sciences call it congestion. To avoid confusion, we refer to the lower branch as hypercongested or jammed traffic.

2 Verhoef (2005) also shows that hypercongestion will occur as a dynamic equilibrium phenomenon, either on a road with a queuing facility in front of its entrance, or on road segments with a downstream bottleneck. This applies provided demand is sufficiently large and that in this case, the average cost curve is not backward-bending, but will eventually rise vertically. As these weaknesses hamper a reasonable economic interpretation of hypercongested traffic states of roads without a downstream bottleneck, deterministic speed-flow models have no longer been considered in recent economic research.

about travel behavior, as our model can be applied directly to road sections for which respective data is available.

A stochastic capacity approach enables us to establish a static model using the speed-flow diagram. 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. 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 idea that the expected costs depend on two possible outcomes, congestion and hypercongestion, has been formalized by Goldmann and Sieg (2019) for the special case of an (experimental) circuit on which density is constant. We augment the existing model to handle real highway traffic data and incorporate the capacity drop.

Small and Chu (2003) argue, that hypercongestion is unstable, occurs as a transient response to demand spikes and should therefore be analyzed with dynamic models. However, Sugiyama et al. (2008), Nakayama et al. (2009) and Tadaki et al. (2013) show that traffic jams can also occur when the number of cars is constant and therefore demand is fixed. Furthermore, our model does not specify a specific traffic state as congested or hypercongested and therefore does not use the average cost curve as supply curve. Because the capacity is stochastic, for the calculation of the expected external costs, it is sufficient to know the probabilities of the two traffic states. At a specific traffic flow q, the probability of hypercongestion (a traffic jam) occuring is p and the probability of congestion (freely flowing traffic) prevailing is $1 - p$. Of course, transitions between traffic states cannot be analyzed in a static model. As we focus on determining expected external costs, the static model does not limit the interpretation of our results.

Each driver only considers his own costs, but not the time losses imposed on other drivers due to increased traffic. We specifically calculate the external costs imposed on other drivers, and we are able to differentiate between congestion costs and hypercongestion costs, where only the former can be calculated in deterministic speed-flow models. We show that the latter, which can be calculated in a stochastic setting, increase total costs.

The remainder of the paper is structured as follows. The following section describes our theoretical model, the third section applies the model to German highway data and the fourth section concludes.

II. Stochastic Speed-Flow Model

Traffic can either be congested at a (high) travel speed of v_h or jammed at a low travel speed of v_l, both depending on the flow q of other cars using the highway during the same time interval. The left side of Figure 1 shows v_h as the upper branch and v_l as the lower branch of the speed-flow function. The probability p that the traffic is jammed also depends on the flow q. The right side of Figure 1 shows a possible distribution function for the probability p of a traffic flow breakdown. For small flows, the traffic breakdown probability is usually small and visually not distinguishable from zero.

We furthermore incorporate the fact that when the traffic flow breaks down, capacity drops (Hall and Agyemang-Duah, 1991; Yuan et al., 2017) from a flow q to a jammed flow of $q_{cd}(q)$. A reduction in capacity means that $0 \leq q_{cd}(q) \leq q$ must hold, and we later use a capacity drop function of $q_{cd}(q) = 0.9\ q$, implying a capacity drop of 10 percent. To summarize, in our model, the travel speed

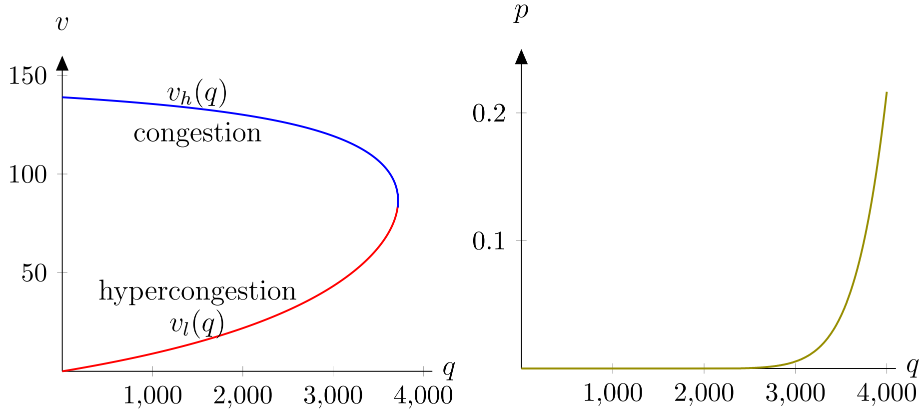


Figure 1. : Stochastic speed-flow model based on Brilon et al. (2005)

depends on q and is either high (but congested) $v_h(q)$ with probability $1 - p(q)$ or low (and jammed) $v_l(q_{cd})$ with probability $p(q)$.

The expected travel speed can be written as

$$E_v(q) = p(q)v_l(q_{cd}(q)) + (1 - p(q))v_h(q). \tag{1}$$

The marginal speed losses that an additional driver imposes on subsequent drivers can be written as

$$\frac{dE_v}{dq} = \frac{dv_h}{dq} - \frac{dp}{dq}\left(v_h(q) - v_l(q_{cd})\right) - p(q)\left(\frac{dv_h}{dq} - \frac{dv_l}{dq_{cd}}\frac{dq_{cd}}{dq}\right). \tag{2}$$

Equation 2 can be split into two parts. The first term is the normal speed loss due to congestion (dv_h/dq), while the second is the hypercongestion adjustment that incorporates the probabilities of a traffic jam. Travel time costs c depend on the speed, which in turn depends on the number of vehicles per hour, and the expected travel time costs C of a driver are

$$C(q) = p(q)c(v_l(q_{cd}(q))) + (1 - p(q))c(v_h(q)). \tag{3}$$

When we assume homogenous drivers, these costs are the average costs of all vehicles. Social costs are $SC = q \cdot C(q)$ and marginal social costs are $MSC = C + q \cdot C'$. In this decision, $q \cdot C'$ is the external effect (on other drivers), which is not taken into account by individual drivers. The marginal external travel time costs are:

$$q\frac{dC}{dq} = q\left[(1-p)\cdot\frac{dc}{dv}\frac{dv_h}{dq} + p\cdot\frac{dc}{dv}\frac{dv_l}{dq_{cd}}\frac{dq_{cd}}{dq} + \frac{dp}{dq}\left(c(v_l(q_{cd})) - c(v_h(q))\right)\right]. \tag{4}$$

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:

$$q\frac{dC}{dq} = q\left[(1-p)\cdot\frac{(-ta)}{v_h^2}\frac{dv_h}{dq} - p\cdot\frac{ta}{v_l^2}\frac{dv_l}{dq_{cd}}\frac{dq_{cd}}{dq} + \frac{dp}{dq}\left(\frac{ta}{v_l} - \frac{ta}{v_h}\right)\right]. \tag{5}$$

Rearranging Equation 5, yields the congestion and hypercongestion costs

(6)
$$q\frac{dC}{dq} = \underbrace{-qta\left[\frac{1}{v_h^2}\frac{dv_h}{dq}\right]}_{\text{Deterministic congestion costs}} \underbrace{-qta\left[p\left(\frac{1}{v_l^2}\frac{dv_l}{dq_{cd}}\frac{dq_{cd}}{dq} - \frac{1}{v_h^2}\frac{dv_h}{dq}\right) + \frac{dp}{dq}\left(\frac{1}{v_h} - \frac{1}{v_l}\right)\right]}_{\text{Stochastic hypercongestion adjustment}}.$$

The first term of Equation 6 represents the congestion costs due to speed losses on the upper branch of the speed-flow curve in a deterministic setting (Deterministic congestion costs). The second term incorporates the probabilities that come into play in a stochastic setting (Stochastic hypercongestion adjustment). As congested traffic only prevails with probability $1 - p$, the congestion effect is overestimated in a deterministic setting, which the term $(p \cdot (.... - 1/v_h^2 \cdot dv_h/dq))$ in the hypercongestion adjustment corrects. The term $(1/v_l^2 \cdot dv_l/dq_{cd} \cdot dq_{cd}/dq)$ in the hypercongestion adjustment displays the marginal costs in hypercongested traffic, and the last term $(dp/dq \cdot (1/v_h - 1/v_l))$ shows the expected speed losses due to phantom traffic jams.

The part labeled stochastic hypercongestion adjustment therefore augments the speed-flow model for hypercongested traffic states and thus incorporates a traffic state that could not have been analyzed with the earlier deterministic speed-flow models. If one wants to exclude the capacity drop, q_{cd} can be set equal to q, which eliminates the corresponding derivative in Equation 6.

Expected average travel time costs are increasing in q if $dC/dq > 0$ which depends on the specific functions that describe the road sector, but is independent of time costs and distance travelled.

III. Application to traffic data of the highway A42

Similar to the U.S. Highway Capacity Manual, the German Highway Capacity Manual (HBS) describes the design capacities of highways and provides standardized methodologies and values for evaluating the performance of highway sections. Underlying research for the HBS comprises amongst others, the specification of the functional forms describing the speed-flow relationships, as well as the functional form of the distribution of the traffic flow breakdown probability. To calculate the effects of Equation 6, we need to know those functional forms. With this knowledge we are able to calculate the expected marginal speed losses depending on the number of cars traversing the highway section. For this reason, we apply the same methodology as in the HBS following Brilon et al. (2005), Brilon and Geistefeldt (2009, 2010); Geistefeldt (2016).[3]

A. Data

We use traffic data for the highway section 44092161 from Straßen.NRW for the highway A42 which is located in the northern Ruhrgebiet in North Rhine-Westphalia. The highway section lies in a metropolitan area and has two lanes and the speed limit is 100 km/h. We employed data for 5-min intervals covering the flow in veh/5 min, speed in km/h and the density in veh/km. Speed and flow

[3] As we want to give an example how to use our theoretical model to calculate congestion costs, we keep the calculation quite simple. We are aware of the fact, that more sophisticated methods are available for example for the calculation of the breakdown probability.

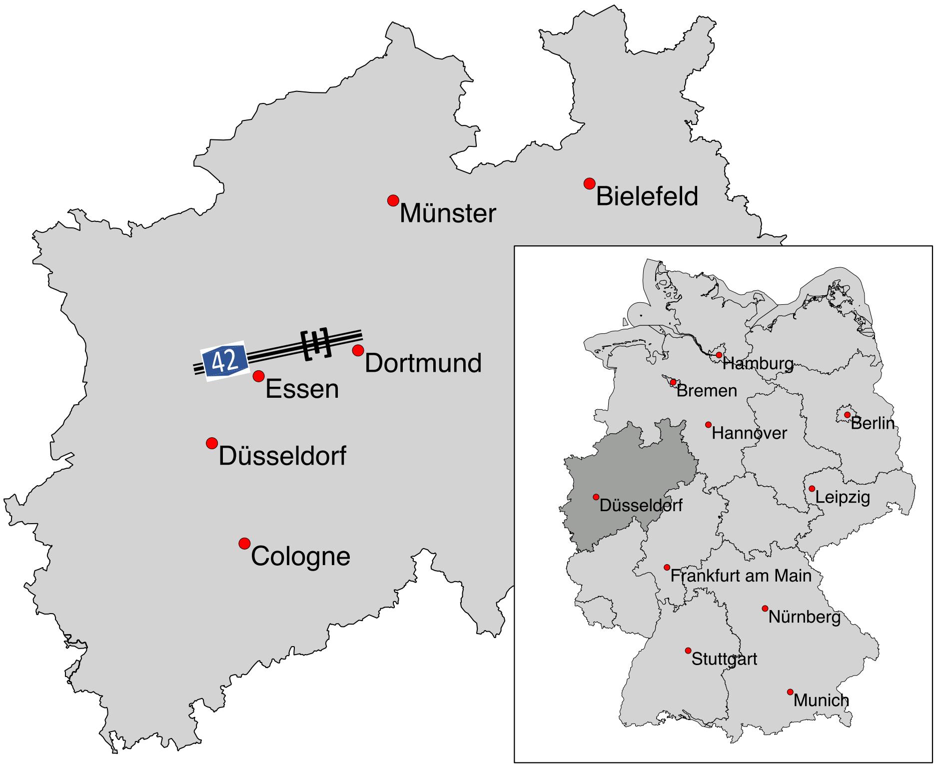


Figure 2. : Location of highway and traffic sensor [|]

are available separately for cars and trucks. Local speeds are converted into space mean speeds following Brilon and Geistefeldt (2010).

As the highway capacity depends on weather conditions and the amount of daylight, we match the traffic data with weather and sunrise and sunset data.[4] By doing so, we can exclude all intervals where road capacity was below the maximum possible capacity. Rain, darkness and frost, for instance, influence road capacity negatively. In addition, this information can also be used directly to analyze the impact of weather conditions on external costs.[5]

B. *Functional forms of speed-flow relationships*

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. Hranac et al. (2006) and Rakha (2009) 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

[4] Data on rainfall and temperatures are from the Deutscher Wetterdienst, Germany's national meteorological service, and sunrise and sunset data are taken from the webpage: `https://www.timeanddate.de/sonne/deutschland/muenster?monat=1&year=2015`.

[5] A figure that shows how rain and darkness increase the marginal external costs is provided in the appendix B.

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 space mean speed, the Van Aerde model can be written as

$$q(v) = \frac{v}{c_1 + c_2/(v_0 - v) + c_3 v}. \tag{7}$$

where c_i are parameters of the function and v_0 is the speed at a flow or density of zero. The Van Aerde function is backward-bending and each value of q can be assigned to one speed of congested v_h, as well as to one speed of hypercongested v_l 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.[6] For this reason, we calibrate the four parameters (c_1, c_2, c_3, v_0) of the Van Aerde function, minimizing the squared errors with respect to speed, flow and density following Brilon and Geistefeldt (2010). The parameters are displayed in Table 1 in the appendix. The model is also employed by Peer et al. (2012) to calculate link-specific free flow travel times.

C. *Probability of traffic flow breakdowns*

In the fundamental diagram, coming from very low traffic flows corresponding to high speeds, the more cars use this highway section per hour, the greater the probability that the traffic will break down (Figure 1). 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; Arnese and Hjelkrem, 2018). Focusing on highway capacity analysis, Brilon et al. (2005) and Brilon and Geistefeldt (2010) use the non-parametric Product Limit Method of Kaplan and Meier (1958) to calculate the breakdown probabilities. The method builds on the idea that for high traffic flows, it is possible to observe either freely flowing/congested traffic or hypercongested traffic in the next interval.[7] For this reason, it is possible to calculate the number of intervals with an observed traffic volume of q which are not followed by traffic breakdowns (censored intervals), and the number of intervals with traffic volume q that are indeed followed by a traffic flow breakdown in the next interval. Inserting this information into the Product Limit Function, enables calculating the breakdown probability. Applying the Product Limit Method to traffic data requires defining a threshold speed, above which traffic is congested/freely flowing, and below which traffic is hypercongested with stop-and-go patterns. Following Brilon et al. (2005) and Tu et al. (2012), we employ a threshold speed of 70 km/h.

Brilon et al. (2005) found that the normal Weibull distribution best fits the non-parametric distribution function of the investigated German motorway sections.

[6] To obtain good estimates for the Van Aerde model, further data cleansing is necessary, such as removing intervals with temporary obstacles. Moreover, all data points where the standard deviation of the speed travelled in the 5-min intervals and in the corresponding 60-min interval exceeded a value of 10, are removed for the estimation of the speed-flow relationship. The data cleansing is described in both Brilon et al. (2005) and Brilon and Geistefeldt (2010).

[7] We only include traffic flow breakdowns at traffic flows greater than 2,400 vehicles per hour. With lesser flows, traffic breakdowns are probably caused by bottlenecks.

The distribution function of the Weibull distribution has the following form:

$$F(q) = 1 - e^{-(q/\beta)^{\alpha}}, \tag{8}$$

where q is the traffic volume, α is the shape parameter and β is the scale parameter. We follow the approach of Brilon et al. (2005) to determine the distribution of traffic flow breakdowns.

As breakdowns of traffic flows occur suddenly, only short time intervals are appropriate for analyzing traffic breakdowns. The time intervals employed in the empirical literature vary between 1-minute (Calvert and Snelder, 2016) and 10-minute intervals (Tu et al., 2012). We follow Brilon et al. (2005) and Brilon and Geistefeldt (2010), who use 5-min intervals to estimate the parameters of the Weibull distribution. By assuming that the variance of the traffic flow is normally distributed over the interval, it is possible to convert the Weibull distribution function to hourly intervals. The procedure is described in detail in Brilon et al. (2005), and is necessary, as the capacity estimation with the Van Aerde function and the calculation of time costs also builds on hourly data. The parameters are presented in Table 1 in the appendix. Figure 3 shows the speed-flow relationship and the Weibull distribution functions. The breakdown probability distribution function for hourly intervals is shifted inwards, as the probability that traffic flow will break down within the next hour is c.p. higher than the probability that the traffic flow will break down within the next five minutes. The displayed distribution functions were calculated including all weather and light conditions. Breakdown probabilities would increase c.p. if only time intervals with, for example, rainfall were considered for calculation.

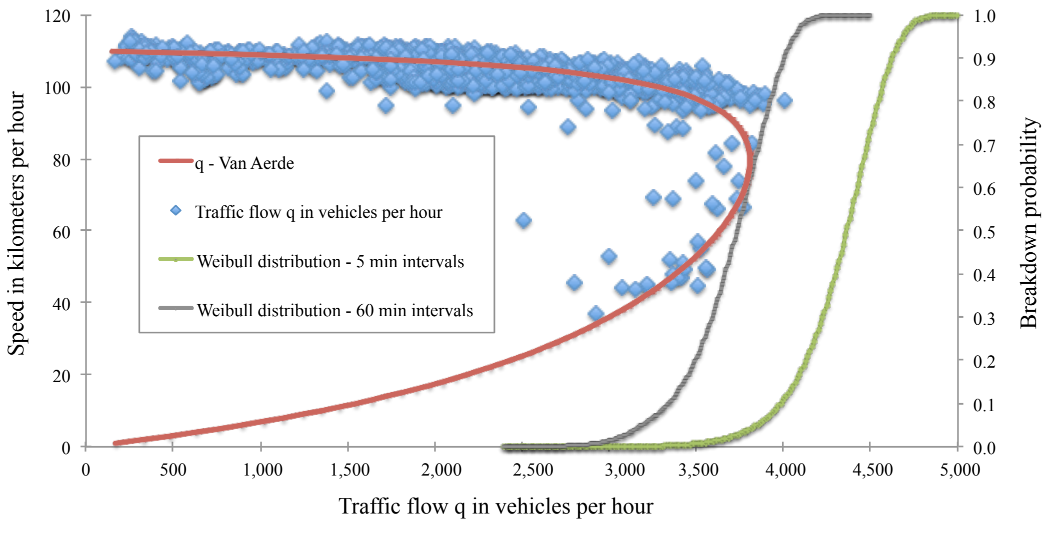


Figure 3. : Van Aerde Function and Weibull Distribution for the A42 (westbound)

Knowing the functional forms of the speed-flow relationships and the breakdown probabilities, we are able to calculate the external costs. The functional form of the Van Aerde function allows for calculating the marginal speed changes caused by additional drivers on the road per hour (from q to $q + 1$). The distribution function of the breakdown probability yields the probability, of the traffic states of congestion or hypercongestion prevailing. More precisely, the breakdown probability measures the probability that the traffic flow will break down in the next interval given the traffic flow in this interval. However, using this probability for the traffic state of hypercongestion is not entirely correct as at this traffic state, flow has already broken down in previous intervals. As the hypercongested intervals are quite short in our data (approximately 14 minutes), the error we make using the breakdown probability also for the state of hypercongestion is small. Furthermore, each additional driver using the road per hour, marginally increases the probability of a breakdown. For this reason, we also need to investigate the changes in p for increases from q to $q + 1$.

D. *Capacity drop*

The capacity drop has received considerable attention in transportation science literature. It describes the observation that the discharge rate of hypercongested traffic is lower than the maximum flow in congested but freely flowing traffic (Hall and Agyemang-Duah, 1991; Yuan et al., 2017). At a traffic breakdown, in our model, the traffic state simply switches from the congested to the hypercongested branch. The capacity drop renders the hypercongested traffic state more persistent, because, due to the lower traffic flow, traffic demand has to fall to a much lower level to dissolve the traffic jam (Treiber and Kesting, 2013).

Research has tended to concentrate on the mechanism of the capacity drop phenomenon at bottlenecks, taking into account various aspects like the impact of driving behavior (Chen et al., 2014), the existence of lane-drops, on-ramps with or without ramp controls (Chung et al., 2007; Srivastava and Geroliminis, 2013) or the impact of different jam types like standing queues or stop-and-go waves, e.g. (Yuan et al., 2017). The results of Yuan et al. (2017) indicate that the outflow of stop-and-go waves is lower than those of standing queues. As stop-and-go waves are especially relevant for phantom jams, their results indicate the importance of this phenomenon in this context.

Estimates of the capacity drop range between 3% and 18%, with Chung et al. (2007) obtaining this entire range of estimates. Hall and Agyemang-Duah (1991) and Srivastava and Geroliminis (2013) have estimates in the medium range of 6% and 15% respectively. The US Highway Capacity Manual (2016) recommends 7% as a default value. Ponzlet (1996) augments the Van Aerde Model by an additional parameter, so as to take the capacity drop into account and applies it to German highway data. He finds that capacity drops by 11% when the traffic flow breaks down.

Our estimates for the capacity drop on the highway section range from 3% eastbound to 13% westbound. Extensive and more sophisticated research on the extent of the capacity drop phenomenon, as performed in this paper, is presented in the above mentioned papers. For this reason, we simply assume a medium value of 10% in our calculations. Within the framework of our model, we assume that the capacity drop only affects hypercongested traffic states. When traffic breaks down, we assume losing 10% of the flow and therefore, the shift from the upper to the lower branch occurs diagonally, resulting in even greater speed losses. We also assume that in the state of hypercongestion, the traffic flow is 10 % lower, so that the dissolution of traffic jams is less efficient.

E. *Travel time costs*

Following the German guidelines for infrastructure planning, 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 (during working hours) and trips of heavy-duty vehicles (trucks). The German methodology handbook for the federal infrastructure plan differentiates between private and business time cost parameters, with both increasing in total trip length (BMVI, 2016, pp. 97-101). The study "Mobility in Germany" contains the average car trip lengths, as well as the trips broken down by purpose (Infas and DLR, 2008, p. 28, p. 89). 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 approximately 18 km is somewhat greater for trips on highways (45 km). This assumption is necessary, because, as mentioned above, the time value function is upward-sloping with the trip length. The corresponding time costs are 8.17 Euro/h. The average length of business trips on highways is assumed to be 100 km (time costs: 30 Euro/h) (BMVI, 2016, pp. 97-101).

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 (BMVI, 2016, pp. 133-134). 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, p. 101). 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 this highway section, 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.

The shares of trips by purpose are also from the Mobility in Germany study, although the trip purposes, including routes on highways may differ from those within urban centers. However, detailed data for highway trips is not available. The same applies to the average rate of vehicle occupancy for cars, to which we apply a value of r_{VO} of 1.1 (BaSt, 2012, p. 8). 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 congestion effect corresponding to the upper branch of the speed-flow curve, barely affects heavy-duty vehicles, as their maximum permissible speed in Germany is 80 kilometers per hour. They do not incur significant travel time prolongation in congested traffic on the upper branch of the speed-flow curve.[8] The travel time cost parameter c_1 is:

$$c_{con} = w_p \cdot r_{VO} \cdot 8.17 \text{ €} + w_b \cdot r_{VO} \cdot 30 \text{ €} = 11.80 \text{ €}, \tag{9}$$

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

The hypercongestion externality is relevant for all vehicles on the highway, including heavy-duty vehicles. For this reason, the weights are somewhat different

[8] As the apex of the Van Aerde function is at a speed of 70 km/h (eastbound) and 80 km/h (westbound), we accept a small error with the assumption that heavy duty vehicles are not affected by travel time prolongation on the upper branch. However, we believe that the error is small and that travel time losses in congested traffic are primarily an issue for cars.

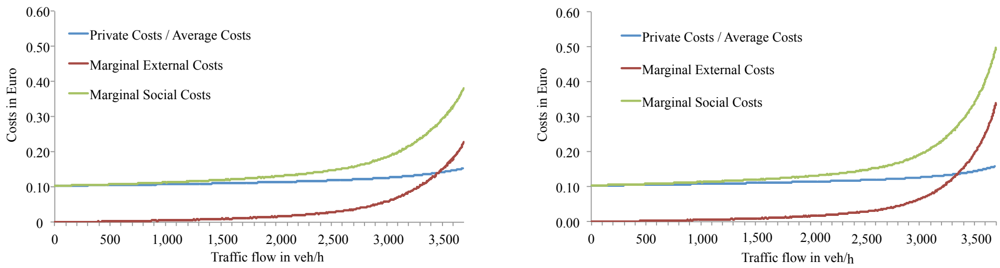
at $w_p = 0.77$, $w_b = 0.10$ and $w_{hd} = 0.13$.

$$c_{hyper} = w_p \cdot r_{VO} \cdot 8.17 \text{ €} + w_b \cdot r_{VO} \cdot 30 \text{ €} + w_{hd} \cdot 30.28 \text{ €} = 14.12 \text{ €} \tag{10}$$

Evaluating the travel time losses due to the normal congestion effect and those due to hypercongestion with the cost parameters, enables us to calculate external congestion costs that depend on the current traffic flow situation.

F. Results

Figure 4 shows the total expected private, marginal and social travel time costs (without and with capacity drop) that have been calculated with the above mentioned time cost parameters.



(a) A42 eastbound (b) A42 eastbound - with CD

Figure 4. : Cost functions

The expected average costs curve (blue line) is upward-sloping. However, compared to the external costs, the slope is quite moderate, underlining the importance of internalizing the external costs in order to obtain socially acceptable quantities.

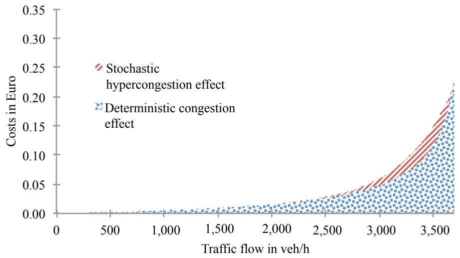
Figure 5 shows the marginal external cost functions for the highway section with and without capacity drop (indicated with CD). The effects are split as in Equation 6 in the deterministic congestion effect (blue - dotted) and the stochastic hypercongestion adjustment (red - striped). As the cost functions display the expected costs of a specific traffic volume q, they increase monotonously in q.

The upward-sloping cost curve and the surge at very high traffic flows are driven by three factors:

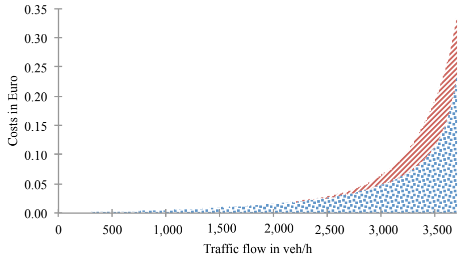
1) The more vehicles that want to use the road at the same time, the more other vehicles are affected by travel time losses. If marginal effects were constant over the entire traffic flow range, this would result in a linearly increasing cost function.

2) However marginal effects of additional drivers are not constant over the entire traffic flow range, as the slope of the Van Aerde function at the apex is much steeper than at low traffic flows.

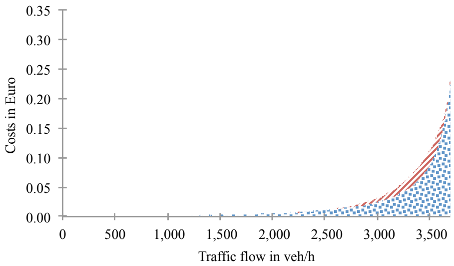
3) The probability that the traffic flow breaks down increases with flow, and therefore, the costs of the shift from congested to hypercongested traffic



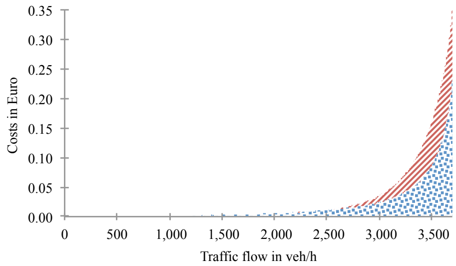
(a) A42 eastbound



(b) A42 eastbound - with CD



(c) A42 westbound

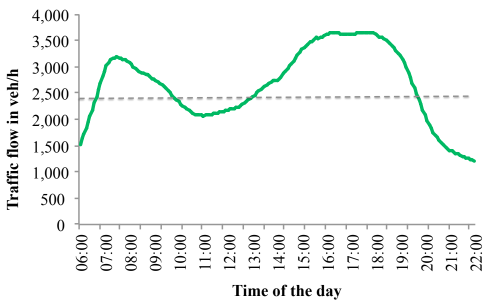


(d) A42 westbound - with CD

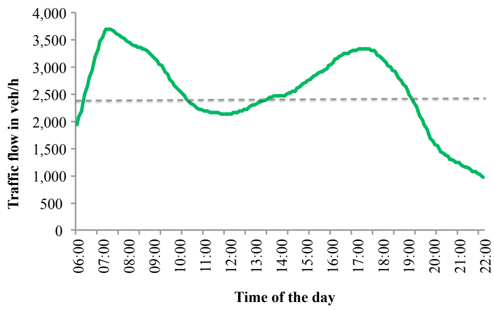
Figure 5. : Marginal external costs without (left) and with (right) capacity drop

become more relevant at higher traffic flows. This overcompensates for the fact that the absolute speed losses due to breakdowns decrease with q.

These marginal cost functions enable us to assign corresponding costs to each traffic flow observed on the highway. Our next step is thus to use the observed traffic flow for an average Thursday (public and school holidays excluded) at this highway section of the A42 (Figure 6).



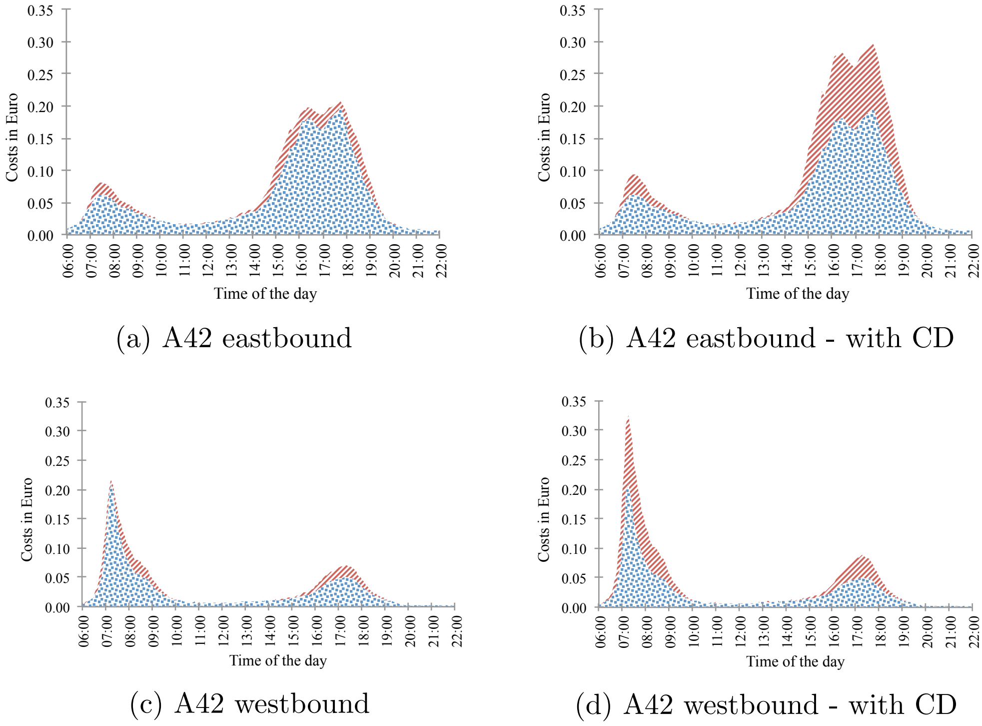
(a) A42 eastbound



(b) A42 westbound

Figure 6. : Traffic flow on an average Thursday

Figure 7 shows the external congestion costs (blue - dotted), as well as the hypercongestion adjustment (red - striped). It is evident that at peak times, due to the increase in probability of a traffic breakdown, the costs of hypercongestion become more pronounced, whereas in off-peak times, these costs equal zero. More precisely, when the flow exceeds approx. 65% of design capacity flow (striped line in Figure 6), the hypercongestion costs start to increase. This effect becomes especially relevant when the capacity drop is taken into account as well (Figure 7 b and d).



(a) A42 eastbound

(b) A42 eastbound - with CD

(c) A42 westbound

(d) A42 westbound - with CD

Figure 7. : External congestion costs (blue) and hypercongestion adjustment (red)

In off-peak periods, the probability of a random breakdown is close to zero, as driver errors, except for those causing accidents and thereby bottlenecks, do not affect the stability of the traffic flow up to a certain saturation level and therefore, the hypercongestion externality is zero as well. Traffic breakdowns at low traffic flows are caused by bottlenecks and should therefore be analyzed with bottleneck models.

In Figure 8 in appendix B we use the information on different weather and daylight conditions to show that this information influences traffic flow conditions and thus also congestion costs. Using only time intervals without daylight and with rainfall, the breakdown probability function shifts inwards and thus the undesired traffic state of hypercongestion becomes more likely. This increases the marginal costs by approximately 7 cents in the afternoon peak on the A42 eastbound compared to Figure 7b. Contrarily, if only favorable weather and light conditions are included in the calculation, external costs would be lower compared to the baseline case.

In total, we identify a currently non-internalized congestion externality for this highway section of a maximum of about 34 cents per vehicle and kilometer. The average externality ranges between 1.4 and 4.0 Euro cents, when costs are spread equally over all intervals. More precisely, the average externality in the eastbound direction lies between 3.3 and 4.0 Euro cents without and with capacity drop respectively. On the westbound section the average externality, ranging within 1.4 and 1.8 Euro cents, is lower due to fewer hours of traffic congestion. The values may be on average large enough to justify a congestion charge, when considering the costs of the charging technology of arguably 2.5 Euro cents per kilometer. If this is not the case yet, as the values of the westbound direction might suggest, decreasing costs of the charging technology and increasing congestion will probably make congestion charges profitable in future.

The absolute size of the external costs we obtained cannot be taken as a reference value for other highway sections, because they are very specific to respective traffic situations. The results depend on the estimated speed-flow relation and the breakdown probability, and these can be very different among different highways and highway sections. Therefore, the extent of congestion and hypercongestion costs are highly specific as well, so that future research should include an analysis of more highway sections with different characteristics regarding the number of lanes or the speed limit, so as to determine which aspects affect external costs. These calculations can also be extended to the entire highway 42 because, for instance, traffic detectors are located every 2.5 kilometers on this highway.

IV. Conclusion

Especially in metropolitan areas, highways are congested during the rush hour. Travel times increase significantly due to congestion, and the resulting additional time and environmental costs place a large burden on economies. There are several possible reasons for congestion at a specific site, but one that is relatively independent of a specific location or road design, is the driving behavior. Driving behavior is, however, the main reason for phantom traffic jams.

Following Sugiyama et al. (2008), we also assume that there are random traffic jam formations that are not caused by bottlenecks. Departing from traffic experiments with stochastic traffic flow breakdowns, we set up a model for calculating their external costs. We show that considering capacity as deterministic ignores parts of the externality, which we refer to as stochastic hypercongestion adjustment.

Directly using the speed-flow data, we calculate external congestion and hypercongestion costs for a German highway section. By incorporating the probabilities in the speed-flow model, we obtain a cost function that increases monotonously with flow and is not backward-bending, as in deterministic speed-flow models. For this reason, a unique cost value can be assigned to each level of traffic flow.

Our results indicate that the stochastic hypercongestion adjustment is not negligible, especially when considering the capacity drop due to traffic flow breakdowns. We show that the costs caused by stochastic traffic flow breakdowns can increase the deterministic congestion costs by up to 50%.

To improve welfare, external costs should be internalized. The congestion and hypercongestion externalities calculated in this paper can only be considered as congestion charges, if demand is inelastic, but in this case congestion charges were redundant and no welfare gains can therefore be expected. Because demand is elastic, for calculating congestion charges, our approach has to be combined with a demand model. Because in our application the cost function is increasing, a unique equilibrium should exist. In this congestion (and hypercongestion) charge equilibrium, congestion externalities will be reduced and our maximum values (of about 0.34 Euro per km) could be considered as an upper bound of the equilibrium congestion charge. But, because users adapt their departure times rather than stopping driving, it can be expected that the peak period will grow. Therefore, in some periods, the equilibrium charge may be higher than our values.

In our model, we also assume that all drivers a equal. Trucks, however, are longer, slower, and heavier, and therefore warrant special treatment and impose higher external congestion costs (Verhoef et al., 1999). Based on Coifman (2015), who finds that many of the critical parameters of the flow-density relationship depend on vehicle length, future research should separate the external effects of trucks and cars on travel times, so as to determine vehicle-type-dependent external costs that could, in a next step, be used to calculate congestion and hypercongestion charges.

Appendices

A. Calibrated parameters

Table 1—: Parameters and values used for application

Parameters of van Aerde Model		
parameter	A42 eastbound	A42 westbound
c_1	0.007521	0.004880
c_2	0.475520	0.092570
c_3	0.000001	0.000163
v_0	114.1	110.0
Weibull-distribution - 60-min intervals, baseline case		
α	13.82	16.55
β	4377	4388
Weibull-distribution - 60-min intervals, rain and darkness		
α	16.07	18.49
β	4162	4331
Travel time cost parameters in Euro[a]		
Effects involving the congested branch	11.80	
Effects involving the hypercongested branch	14.12	

[a] Own calculations based on Infas and DLR (2008) and BMVI (2016).

B. Additional costs caused by rain and darkness

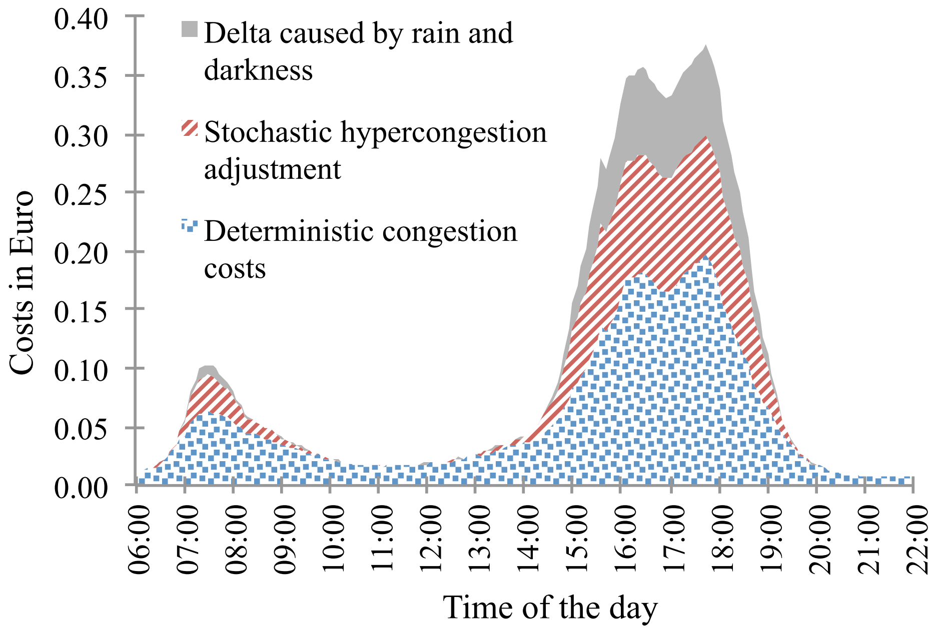


Figure 8. : A42-eastbound: Additional costs depending on outside conditions

REFERENCES

Arnese, P. and Hjelkrem, O. A. (2018). An estimator for traffic breakdown probability based on classification of transitional breakdown events. *Transportation Science*, 52(3):593–602.

Arnott, R. (2013). A bathtub model of downtown traffic congestion. *Journal of Urban Economics*, 76(1):110–121.

Arnott, R., de Palma, A., and Lindsey, R. (1990). Economics of a bottleneck. *Journal of Urban Economics*, 27:111–130.

Arnott, R., de Palma, A., and Lindsey, R. (1993). A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *American Economic Review*, 83(1):161–179.

Arnott, R., Kokoza, A., and Naji, M. (2016). Equilibrium traffic dynamics in a bathtub model: A special case. *Economics of Transportation*, 7-8:38–52.

BaSt (2012). *Zeitbewertung in ökonomischen Planungsrechnungen*. Bundesanstalt für Straßenwesen.

BMVI (2016). *Methodenhandbuch zum Bundesverkehrswegeplan 2030*. Federal Ministry of Transport and Digital Infrastructure.

Brilon, W. and Geistefeldt, J. (2009). Implications of the random capacity concept for freeways. In *Proceedings of the International Symposium on Freeway and Tollway Operations*.

Brilon, W. and Geistefeldt, J. (2010). Überprüfung der Bemessungswerte der HBS für Autobahnabschnitte außerhalb der Knotenpunkte. *Forschung Straßenbau und Straßenverkehrstechnik*.

Brilon, W., Geistefeldt, J., and Regler, M. (2005). Reliability of freeway traffic flow: A stochastic concept of capacity. In *Prooceedings of the 16th International Symposium on Transportation and Traffic Theory*, pages 125–144.

Brilon, W., Geistefeldt, J., and Zurlinden, H. (2007). Implementing the concept of reliability for highway capacity analysis. *Transportation Research Record*, pages 1–8.

Calvert, S. C. and Snelder, M. (2016). Influence of weather on traffic flow: an extensive stochastic multi-effect capacity and demand analysis. *EuropeanTransport Trasporti Europei*, 60.

Chen, D., Ahn, S., Laval, J., and Zheng, Z. (2014). On the periodicity of traffic oscillations and capacity drop: The role of driver characteristics. *Transportation Research Part B*, 59:117–136.

Chung, K., Rudjanakanoknad, J., and Cassidy, M. J. (2007). Relation between traffic density and capacity drop at three freeway bottlenecks. *Transportation Research Part B: Methodological*, 41:82–95.

Coifman, B. (2015). Empirical flow-density and speed-spacing relationships: Evidence of vehicle length dependency. *Transportation Research Part B: Methodological*, 78:54–65.

Elefteriadou, L., Roess, R. P., and McShane, W. R. (1995). Probabilistic nature of breakdown at freeway merge junctions. *Transportation Research Record*, 1484:80–89.

Fosgerau, M. (2015). Congestion in the bathtub. *Economics of Transportation*, 4(4):241–255.

Fosgerau, M. and Small, K. A. (2013). Hypercongestion in downtown metropolis. *Journal of Urban Economics*, 76(1):122–134.

Geistefeldt, J. (2016). Assessment of basic freeway segments in the German Highway Capacity Manual HBS 2015 and beyond. *Transportation Research Procedia*, 15:417–425.

Goldmann, K. and Sieg, G. (2019). Economic implications of phantom traffic jams: evidence from traffic experiments. *Transportation Letters: The International Journal of Transportation Research*, doi.org/10.1080/19427867.2019.1611077.

Hall, F. L. and Agyemang-Duah, K. (1991). Freeway capacity drop and the definition of capacity. *Transportation Research Record*, pages 91–98.

He, X., Chen, X., Xiong, C., Zhu, Z., and Zhang, L. (2016). Optimal time-varying pricing for toll roads under multiple objectives: A simulation-based opimization approach. *Transportation Science*, 51(2):412–426.

Highway Capacity Manual (2016). *Highway Capacity Manual: A Guide for Multimodal Mobility Analysis*. Transportation Research Board.

Hranac, R., Sterzin, E., Krechmer, D., Rakha, H., and Farzaneh, M. (2006). Empirical studies on traffic flow in inclement weather. Technical report, U.S Department of Transportation - Federal Highway Administration.

Infas and DLR (2008). *Mobilität in Deutschland 2008, Ergebnisbericht: Struktur - Aufkommen - Emissionen - Trends.* https://www.bmvi.de/SharedDocs/DE/Anlage/VerkehrUndMobilitaet/mid-2008-ergebnisbericht.pdf.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.

Lorenz, M. and Elefteriadou, L. (2000). A probabilistic approach to defining freeway capacity and breakdown. *Transportation Research Circular E-C018: 4th International Symposium on Highway Capacity.*

Nakayama, A., Fukui, M., Kikuchi, M., Hasebe, K., Nishinari, K., Sugiyama, Y., Tadaki, S., and Yukawa, S. (2009). Metastability in the formation of an experimental traffic jam. *New Journal of Physics*, 11(083025).

Peer, S., Koopmans, C. C., and Verhoef, E. T. (2012). Prediction of travel time variability for cost-benefit analysis. *Transportation Research Part A*, 46(79-90).

Ponzlet, M. (1996). *Auswirkungen von systematischen und umfeldbedingten Schwankungen des Geschwindigkeitsverhaltens und deren Beschreibung in Verkehrsflussmodellen.* PhD thesis, Lehrstuhl für Verkehrswesen Ruhr-Universität Bochum.

Rakha, H. (2009). Validation of van Aerde's simplified steady-state car-following and traffic stream model. *Transportation Letters: The International Journal of Transportation Research*, 1:225–242.

Schönhof, M. and Helbing, D. (2007). Empirical features of congested traffic states and their implications for traffic modeling. *Transportation Science*, 41(2):135–166.

Small, K. A. (2015). The bottleneck model: an assessment and interpretation. *Economics of Transportation*, 4(1–2):110–117.

Small, K. A. and Chu, X. (2003). Hypercongestion. *Journal of Transport Economics and Policy*, 37(3):319–352.

Srivastava, A. and Geroliminis, N. (2013). Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. *Transportation Research Part C*, 30:161–177.

Sugiyama, Y., Fukui, M., Kikuchi, M., Hasebe, K., Nakayama, A., Nishinari, K., Tadaki, S.-i., and Yukawa, S. (2008). Traffic jams without bottlenecks — experimental evidence for the physical mechanism of the formation of a jam. *New Journal of Physics*, 10(033001).

Tadaki, S., Kikuchi, M., Fukui, M., Nakayama, A., Nishinari, K., Shibata, A., Sugiyama, Y., Yosida, T., and Yukawa, S. (2013). Phase transition in traffic jam experiments on a circuit. *New Journal of Physics*, 15(103034).

Taylor, N. B., Bourne, N., Notley, S. O., and Skrobanski, G. (2008). Evidence for speed flow relationships. In *(Proceedings) European Transport Conference 2008*, pages 1–19.

Treiber, M. and Kesting, A. (2013). *Traffic Flow Dynamics*. Springer.

Tu, H., Li, H., van Lint, H., and van Zuylen, H. (2012). Modeling travel time reliability of freeways using risk assessment techniques. *Transportation Research Part A*, 46:1528–1540.

van Aerde, M. (1995). A single regime speed-flow-density relationship for freeways and arterials. *Proceedings of the 74th Transportation Research Research Board Annual Meeting, Washington D.C.*, pages 1–8.

Verhoef, E. T. (1999). Time, speeds, flows and densities in static models of road traffic congestion and congestion pricing. *Regional Science and Urban Economics*, 29(3):341–369.

Verhoef, E. T. (2005). Speed-flow relations and cost functions for congested traffic theory and empirical analysis. *Transportation Research Part A*, 39:792–812.

Verhoef, E. T., Rouwendal, J., and Rietveld, P. (1999). Congestion caused by speed differences. *Journal of Urban Economics*, 45(3):533–556.

Vickrey, W. S. (1969). Congestion theory and transport investment. *American Economic Review*, 59:251–261.

Xiao, L.-L., Huang, H.-J., and Liu, R. (2015). Congestion behavior and tolls in a bottleneck model with stochastic capacity. *Transportation Science*, 49(1):46–65.

Yuan, K., Knoop, V. L., Leclercq, L., and Hoogendoorn, S. P. (2017). Capacity drop: a comparison between stop-and-go wave and standing queue at lane-drop bottleneck. *Transportmetrica B: Transport Dynamics*, 5:2:145–158.

Westfälische Wilhelms-Universität Münster, Institute of Transport Economics, Working Paper Series

21. **"Costs and benefits of a bicycle helmet law for Germany"**
    by Gernot Sieg, March 2014

22. **"Time-Declining risk-adjusted social discount rates for transport infrastructure planning"**
    by Kathrin Goldmann, April 2017

23. **"To pay or not to pay for parking at shopping malls - A rationale from the perspective of two-sided markets"**
    by Inga Molenda and Gernot Sieg, November 2017

24. **"A duopoly of transportation network companies and traditional radio-taxi dispatch service agencies"**
    by Thorsten Heilker and Gernot Sieg, November 2017

25. **"Airport Efficiency in Pakistan – A Data Envelopment Analysis with Weight Restrictions"**
    by David Ennen and Irem Batool, November 2017

26. **"Economic implications of phantom traffic jams: Evidence from traffic experiments"**
    by Kathrin Goldmann and Gernot Sieg, December 2018

27. **"Evaluating the Transport-Mode-Specific Trade Effects of Different Transport Infrastructure Types"**
    by Jan Wessel, February 2019

28. **"The impact of delays on the welfare effects of on-track competition: The case of transfer passengers with operator-tied tickets"**
    by Christina Brand and Gernot Sieg, December 2019

29. **"Ride-Hailing Services in Germany: Potential Impacts on Public Transport, Motorized Traffic, and Social Welfare"**
    by David Ennen and Thorsten Heilker, January 2020

30. **"Quantifying the phantom jam externality: The case of an Autobahn section in Germany"**
    by Kathrin Goldmann and Gernot Sieg, March 2020

For a complete list of Working Papers published by Westfälische Wilhelms-Universität Münster, Institute of Transport Economics, please visit the website (http://www.iv-muenster.de)