An Empirical Macroscopic Bottleneck Model for the Randstad Area

Erik Verhoef*[1], Bas Turpijn[2], Maarten Verhoef[1]

[1] Vrije Universiteit Amsterdam

[2] Rijkswaterstaat

SHORT SUMMARY

We propose an area-wide Macroscopic Bottleneck Model for the Dutch Randstad area and use it to estimate a relation between dynamic aggregate inflow patterns and travel delays on highways. We find plausible parameter estimates, and a strong time variation of marginal external congestion costs, which are €9.68 in the busiest 30 minutes and only €2.08 for the other 60 minutes in the busiest 90 minutes.

Keywords: Traffic congestion, Congestion modelling, Bottleneck model, Macroscopic Fundamental Diagram, Traffic Flow Theory.

1. INTRODUCTION

Recent years have witnessed an upsurge in studies that apply macroscopic, single-facility models to represent congestion in urban areas. Coined "Macroscopic Fundamental Diagram" (MFD) models, these models relate instantaneous measures of (space-)averaged speed and flow in urban areas, producing relations that mimic patterns found with the standard fundamental diagram that relates instantaneous and local measures of speed and density on a single facility; see, for example, Geroliminis and Dagonzo (2007, 2008); Amirgholy and Gao (2017); and Mariotte et al. (2017). Closely related to these are economic dynamic equilibrium models that have become known as "bathtub models", which describe traffic dynamics in a single-facility representing a downtown area, where traffic conditions are spatially homogeneous but vary in continuous time (Arnott, 2013; Fosgerau, 2015).

These models have in common that the traffic volume measure used as the argument in the function describing speed is traffic flow observed *within* the facility modelled. Plausible as this may seem at first sight, an alternative measure becomes relevant when for traffic conditions close to the capacity of the facility, queues may be forming upstream of that facility, consisting of vehicles waiting to enter. Alternative model formulations can then seek to relate speeds or travel times to either traffic flow within the facility of interest, or to the arrival rate of vehicles at the back of the queue(s) (where "arrivals at the back of the queue" are somewhat confusingly also frequently referred to as "departures from home"). There is not such a thing as a single "correct" measure for vehicles per unit of time in such a setting: both measures are relevant. Their instantaneous values may however diverge considerably over clock time, and for many plausible model specifications the flow inside the facility is endogenous, depending on the history of arrival rates of vehicles at the facility's entrance(s).

The temporal divergence of instantaneous flow inside the facility and arrival rate at its entrance causes an important divergence in the shape of the relation between vehicles per unit of time on the one hand, and travel times experienced on the other; an important basis for economic equilibrium models of congestion. Notably, as shown both for an empirical bottleneck and for a theoretical dynamic model based on car-following theory in Verhoef (2001, 2003, 2005), for a facility for which instantaneous and local traffic conditions are characterized by the conventional backward-bending speed-flow relation, the relation between steady-state arrival rates and per-user travel times may still be strictly monotonously increasing, also when trips experience queuing at hypercongested speeds.

These observations motivated us to develop an empirical macroscopic (single-facility) model of congestion for an urban area, in which the measure of volume is not the endogenous (on past inflows) space-averaged flow in the facility as in MFD and bathtub models, but rather the inflow into the facility. We coin our model a "Macroscopic Bottleneck", to reflect an important similarity with the economic bottleneck model of traffic congestion as developed by Vickrey (1969) and Arnott, de Palma and Lindsey (1993): the use of the arrival rate of "new" users at the entrance as the argument in the dynamic travel time function. More specifically, we investigate how travel delays are determined by aggregate inflow over the peak as well as its dynamic pattern. This is the feauture of the model that makes it consistent with the bottleneck model. The findings are directly relevant for congestion policies, as these show how a change in overall traffic volume and its time pattern would affect congestion levels.

We emphasize that the aggregate nature of the macroscopic model makes it also differ from the conventional bottleneck model. First, whereas the latter typically assumes a given constant capacity, the fact that the area we study has many, spatially dispersed exits, which in reality are used with different time profiles without any reason to expect that all exits should operate at capacity at the same clock times, makes this single-capacity measure having no direct equivalent in our model. Moreover, the true bottlenecks in the network we study will often be between points of entry and exit of the network, so that unlike the assumption in the standard bottleneck model, the exits are not the effective capacity constraint. As we can relate travel delays directly to inflow patterns only, we also do not need to specify exit capacity for our study, but it is important to highlight this difference with the standard bottleneck model: we do not assume nor specify any maximum capacity, even though we find a range of observations where our estimated inflow - travel delay relation becomes nearly vertical. A second distinction is that we need to account for phase differences between different points of inflows; put simply: closer to an economic centre, the moment of inflow that corresponds with an arrival at the most desirable arrival moment will be later than for a point of entry that is further away. This we tackle by allowing those moments to be different for the entry points we consider. Therefore, the parallels with the textbook bottleneck model are strong, notably in the focus on inflows rather than flows, but not complete.

2. METHODOLOGY

Study area and data

We apply our model to the entire Dutch Randstad Area; the urbanized area in the western part of the country that contains the cities of Amsterdam, Rotterdam, The Hague and Utrecht. The facility considered consists of the highway network in that area and close to it, since we measure inflows into the highway system also for entries close to but outside the Randstad. The map in Figure 1 reflects that the vast majority of highway travel delays on Dutch national highways occurs in the area considered.

To estimate our model, we use data over the years 2019-2020, thus benefitting in empirical identification from the variation that COVID-19 lockdown measures have brought to commuting behaviour.

Aggregate travel delays are measured as the daily total vehicle hours lost on Dutch national highways, as computed by Rijkswaterstaat, the public agency monitoring traffic conditions on Dutch national highways. We explain these daily travel delays from traffic volumes in the morning and afternoon peaks on weekdays. A strong correlation naturally exists between these ("what goes out to work in the morning, must come home in the evening"), which easily causes problems of collinearity in estimation. We therefore add up traffic volumes in the morning and afternoon peaks, but do include a daily measure for imbalance between the two (*Imbalance*), to reflect that non-linearity of travel delay functions would imply that the same daily total distributed less evenly would lead to higher aggregate travel delays.

Figure 1: Study area (inside blue dashed line) and travel delays on Dutch national highways (May 2019-April 2020)

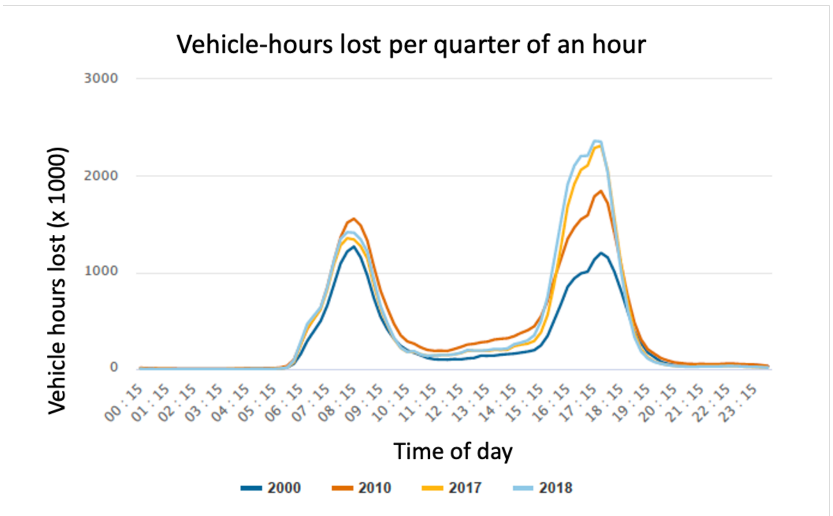

Figure 2: Yearly vehicle hours lost by clock time on Dutch national highways (pre COVID-19)

Figure 2 show the general time patterns of travel delays on Dutch highways in the years before COVID-19. On the basis of these patterns, we define as the total periods two six-hour blocks: 06:00 – 12:00 in the morning, and 14:00 – 20:00 in the afternoon. For these time windows we use quarterly inflows into the highway network on 345 entry points inside the study area indicated with the blue dashed line in Figure 1.

To capture the effect of the time differentiation of inflows on total travel delays we subdivide both the morning and evening period into three parts, following the labelling illustrated in Figure 3.

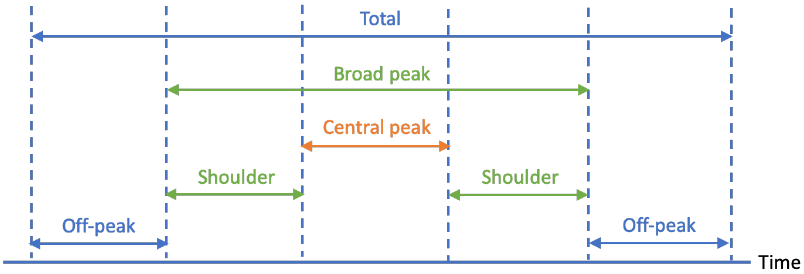

Figure 3: Labelling of time windows

The “Total” periods consist, for each observation point, of the 6-hour windows just mentioned. Within each six-hour period, we define for each entry point individually, and for each working day and for morning and evening separately, the “Broad Peak” as the consecutive period of Y quarters that, among all possible Y-quarter consecutive periods, contains the highest number of vehicles passing. Similarly, we define the “Central Peak” as the busiest consecutive X quarters. The Central Peak thus falls within the Broad Peak; for now we have set X at 2 quarters, and Y at 6 quarters. The part of the Broad Peak that is not in the Central Peak is called the “Shoulder” and it thus contains 4 quarters. The parts of the Total Period that are not in the Broad Peak is called the “Off-peak”.

For each working day, we then add up, over all 345 entry points and over the morning and evening period, the numbers of trips in each of these periods just distinguished, to obtain our daily quantity measures: *Central Peak Trips*, *Broad Peak Trips*, *Total Trips*; as well as *Shoulder Trips* and *Off-peak Trips*. This leads to traffic flows (per quarter of an hour) that are, on average (over the working days), 76 633 in the Central Peak; 62 542 in the Shoulder; and 44 539 in the Off-peak. In the empirical model we will use numbers of trips, so the product of these flows and the length of the period. *Imbalance* is defined over Total Trips as the absolute value of the difference between morning trips over daily trips and 0.5 (so that it increases when either morning or evening traffic is heavier).

For our estimations we use as the dependent variable the travel delays per trip. This is to make the relations resemble average travel time functions, such as the BPR function. Since we only observe delays, free-flow travel times do not enter our model; which is fine given our purpose. The average delay is, for each working day, the total vehicle hours lost divided by the total number of trips. When a given individual travels in the morning and evening period and incurs average delays, (s)he will thus incur that average delay twice per day.

Figure 4 (left panel) shows the scatter diagram relating daily total trips to this average delay. In other words, if indeed the vast majority of travel delays on Dutch highways is incurred by travelers entering via one of the 345 entry points we consider, they will suffer, per entry, from an average time loss that may range up to 0.1 hours or six minutes. The diagram displays a strong relation between these two variables, which in itself is not too surprising, but reinforcing our faith in the plausibility of the assumption that our demarcation of the study area considered makes good sense. It is also clear that the relation is highly non-linear. For that reason, we will be using the natural logarithm of the average lost hours per user in our estimations. The resulting relation, in the right panel, indeed appears close to linear.

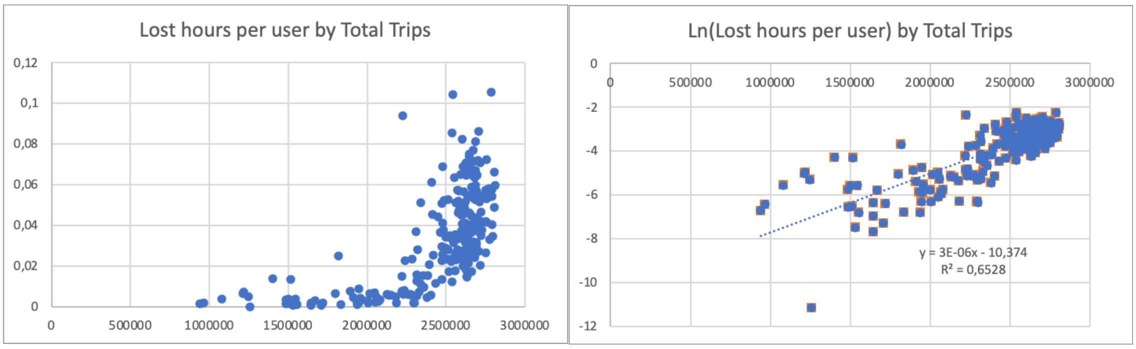

Figure 4: Average per-user delays (left panel) and its natural logarithm (right panel) by Total Trips

Finally, we use a number of covariates to capture the possible impact of external conditions on the relation between traffic volumes and travel delays, all defined such that we expect a positive coefficient. These are:

- *Darkness*: 1 minus the share of the day that the sun is up;
- *Rain*: the summation of millimeters of precipitation in three main weather stations in the area considered: Rotterdam, Schiphol Airport and De Bilt (in the more inland province of Utrecht);
- *Freeze*: a dummy equaling 1 if the minimum temperature is below 1 degree Celcius.

An observation, and a dot in Figure 4, is thus a single working day. The estimations in this extended abstract cover 11 months in the years 2019-2020: Jan, Feb, Mar, May, June; and Dec 2019. For final model estimates more months will be added.

3. RESULTS AND DISCUSSION

Table 1 presents the estimation results, including some robustness checks. The latter are needed especially also because there is a strong correlation between traffic volumes in the different sub-periods. The robustness checks confirm the main findings from the preferred model (for which 5 outliers with a standardized residual exceeding 2.5 have been removed compared to the model in the second column).

The impact of an additional vehicle in the Central Peak is around 5 times as large as the point estimate for vehicles in the Shoulder. The latter is not even statistically significant, suggesting that only vehicles entering in the busiest 30 minutes are by far most decisive for total travel delays. The number of Off-peak trips is also statistically insignificant, and even has the wrong sign in the preferred model. The coefficient for the Central Peak trips is by far the most robust in our sensitivity checks, confirming these conclusions. The increase in parameter estimate and statistical significance in the robustness checks for Shoulder and Off-peak Trips confirms the correlation with Central Peak Trips. Collinearity affects the efficiency but not the unbiasedness of parameter estimates, so that the parameters in the preferred model represent our point estimates for Shoulder and Off-peak trips. The covariates have the expected sign, but are not statistically significant (at the 0.05 level).

Table 1: Estimation Results

	Preferred model	*Preferred model with outliers*	*Robustness check Central Peak*	*Robustness check Shoulder*	*Robustness check Off-peak*
Intercept	-12.05 (2,76 E-91)	-12.12 (2.77 E-82)	-12.22 (3.05 E-89)	-12.21 (4.45 E-88)	-11.91 (1.17 E-68)
Darkness	0.116 (0.698)	0.106 (0.764)	0.200 (0.548)	0.114 (0.736)	1.122 (0.004)
Freeze	0.158 (0.073)	0.189 (0.0700)	0.195 (0.057)	0.217 (0.037)	0.270 (0.025)
Rain	0.00153 (0.102)	0.00165 (0.136)	0.00169 (0.123)	0.00147 (0.186)	0.00211 (0.109)
Imbalance	7.922 (1.95 E-07)	7.313 (3.53 E-05)	7.406 (6.70 E-06)	8.521 (6.04 E-07)	8.055 (0.00011)
CentralPeakTrips	1.946 E-05 (0.037)	1.898 E-05 (0.086)	2.561 E-05 (1.93 E-59)		
ShoulderTrips	4,182 E-06 (0.521)	5.586 E-06 (0.468)		1.568 E-0.5 (6.04 E-07)	
OffPeakTrips	-2,146 E-07 (0.694)	-5.064 E-07 (0.426)			4.38 E-06 (2.05 E-41)
R^2	0.805	0.765	0.765	0.760	0.662
Obs	231	236	236	236	236

P-value between brackets

Using the point estimates from our preferred model, we can construct marginal external cost functions for Central Peak Trips (CPT), and for Shoulder Peak Trips (ST). These are shown in Figure 5. We assume a value of time of 10 Euro/hr, and express the MEC as the cost imposed on all other travelers, in the Total Period, from an additional vehicle in the period considered. It thus also tells us by how much travel cost for all other drivers change if we move one vehicle from the one period to the other. Not surprisingly, the marginal external cost is much higher in the Central Peak. It may reach values as high as €30 for the busiest days in our data, and obtains at the sample average a value of €9.68. For Shoulder Trips, the MEC at the sample average is €2.08.

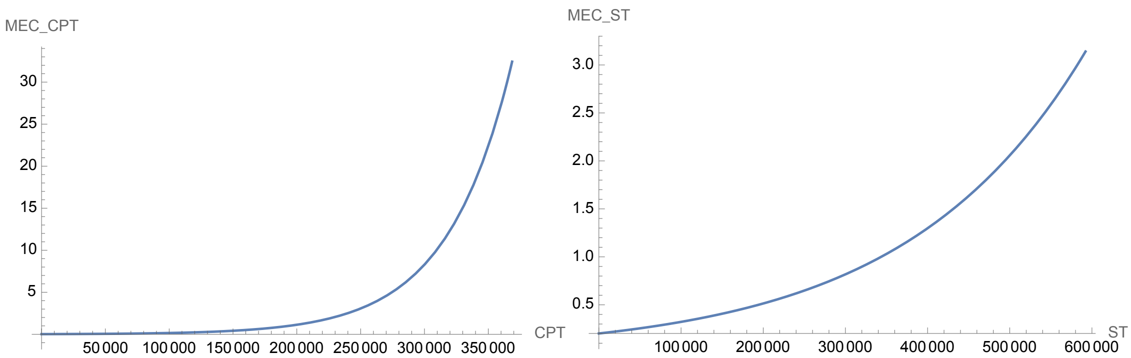

Figure 5: Marginal external cost for Central Peak Trips (left panel) and Shoulder Trips (right panel)

We emphasize that these should not be misinterpreted as optimal access tolls, even though the interpretation is close to it. One caveat is that first-best tolls are equal to marginal external costs *in the optimum*. If a central peak toll would reduce Central Peak trips to 90% (80%) of the sample average, the marginal external cost drops to 5.25 Euro (2.82 Euro). Second, the macroscopic nature of our model equalizes all drivers in all respects. In reality, an individual's marginal external cost from an entry depends at least on the place of entry and exit, ignoring other sources of heterogeneity. The second-best uniform access toll will be a weighted average of the resulting distribution of marginal external costs, and is likely different from the values we now show, even though the order of magnitude will be indicative.

4. CONCLUSION

We propose an area-wide Macroscopic Bottleneck Model for the Dutch Randstad area and use it to estimate a relation between dynamic aggregate inflow patterns and travel delays. We find plausible parameter estimates, judged by their relative sizes and by the absolute values of marginal external costs. We find a strong time variation of marginal external congestion costs, which are €9.68 in the busiest 30 minutes and only €2.08 for the other 60 minutes in the busiest 90 minutes.

REFERENCES

Amirgholy, M. and H.O. Gao (2017) "Modeling the dynamics of congestion in large urban networks using the macroscopic fundamental diagram: User equilibrium, system optimum, and pricing strategies" *Transportation Research Part B: Methodological* **104** 215–237. doi:10.1016/j.trb.2017.07.006.

Arnott, R. (2013) "A bathtub model of downtown traffic congestion" *Journal of Urban Economics* **76** 110–121. doi:10.1016/j.jue.2013.01.001.

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

Fosgerau, M. (2015) "Congestion in the bathtub" Economics of Transportation **4** 241–255. doi:10.1016/j.ecotra.2015.08.001.

Geroliminis, N. and C.F. Daganzo (2008) "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings" *Transportation Research Part B: Methodological* **42** 759–770. doi:10.1016/J.TRB.2008.02.002.

Geroliminis, N. and C.F. Daganzo (2007) "Macroscopic modeling of traffic in cities" *TRB 86th Annual Meetings* 07–0413. doi:10.1002/tcr.201100032.

Mariotte, G., L. Leclercq, L. and J.A. Laval (2017) "Macroscopic urban dynamics: Analytical and numerical comparisons of existing models" *Transportation Research Part B: Methodological* **101** 245–267. doi:10.1016/j.trb.2017.04.002.

Verhoef, E.T. (2001) "An integrated dynamic model of road traffic congestion based on simple car-following theory: exploring hypercongestion" Journal of Urban Economics **49** 505-542. doi:10.1006/JUEC.2000.2203.

Verhoef, E.T. (2003) "Inside the queue: hypercongestion and road pricing in a continuous time – continuous place model of traffic congestion" Journal of Urban Economics **54** 531-565. doi:10.1016/S0094-1190(03)00078-0.

Verhoef, E.T. (2005) "Speed-flow relations and cost functions for congested traffic: theory and empirical analysis" Transportation Research **39A** (7-9) 792-812. doi:10.1016/j.tra.2005.02.023.

Vickrey, W.S. (1969) "Congestion theory and transport investment" *American Economic Review (Papers Proceedings)* **59** 251–260.