Predicting Travel Time Variability for Cost–Benefit Analysis

Stefanie Peer[1,2]
Carl Koopmans[1]
Erik T. Verhoef[1,2]

[1] VU University Amsterdam;
[2] Tinbergen Institute.

Predicting Travel Time Variability for Cost-Benefit Analysis

Stefanie Peer *
VU Amsterdam
Tinbergen Institute Amsterdam/Rotterdam

Carl Koopmans
VU Amsterdam

Erik T. Verhoef
VU Amsterdam
Tinbergen Institute Amsterdam/Rotterdam

Abstract

Unreliable travel times cause substantial costs to travelers. Nevertheless, they are not taken into account in many cost-benefit-analyses (CBA), or only in very rough ways. This paper aims at providing simple rules on how variability can be predicted, based on travel time data from Dutch highways. The paper uses two different concepts of travel time variability. They differ in their assumptions on information availability to drivers. The first measure is based on the assumption that, for a given road link and given time of the day, the expected travel time is constant across all working days (rough information: RI). In the second case, expected travel times are assumed to reflect day-specific factors such as weather conditions or weekdays (fine information: FI). For both definitions of variability, we find that the mean travel time is a good predictor of variability. On average, longer delays are associated with higher variability. However, the derivative of travel time variability with respect to delays is decreasing in delays. It can be shown that this result relates to differences in the relative shares of observed traffic 'regimes' (free-flow, congested, hyper-congested) in the mean delay. For most CBAs, no information on the relative shares of the traffic regimes is available. A non-linear model based on mean travel times can be used as an approximation.

---

*Corresponding Author: VU University, Department of Spatial Economics, De Boelelaan 1105, 1081 HV, Amsterdam, The Netherlands, E-mail: `speer@feweb.vu.nl`

1   Introduction

Travel times on roads are usually not stable over time. Variations occur in demand for travel as well as in road capacity, and these cause travel times to vary. Parts of these fluctuations are known to travelers. Especially frequent travelers are aware of recurrent congestion patterns during peak hours, or the adverse effect of heavy rain on travel times. This paper focuses on variations in travel times that are not expected by drivers. They cause drivers to arrive at their destinations earlier or later than expected. In most cases, such schedule delays come at a cost to drivers. They might face waiting times, or the need to reschedule activities. With most people being risk-averse, the uncertainties in respect to arrival time might be accompanied by feelings of stress and anxiety.

To include the costs associated with unreliable travel times in cost-benefit analysis (CBA), both the drivers' (monetary) valuation of unreliable travel times and the extent of unreliability need to be known. While the values drivers attach to travel time variability and scheduling delays have been derived in various stated and revealed preference experiments (e.g. Hensher, 2001; Lam and Small, 2001; Small, 1982), only little research has been done on explaining and forecasting the extent of travel time variability. Most of the research in this area is based on simulation studies. The outcomes of these studies are usually not very well transferable to CBA, as they focus on stylized cases of traffic networks without the use of empirical data (e.g. Li, Bliemer, and Bovy, 2009; Nagel and Rasmussen, 1995; Nicholson and Du, 1997). Studies that explicitly predict the variability of travel times based on empirical data, and calculate the resulting costs , were done by Eliasson (2006) and Fosgerau, Hjorth, Brems, and Fukuda (2008). They both point out the importance of including variability-related costs into CBA, suggesting that 10–15% of the costs associated with changes in travel times can be attributed to changes in variability.

In order to predict travel time variability, its main determinants need to be known. A common result of previous research is that the main explanatory factor of travel time variability is mean travel time. Fosgerau (2010) uses Jensen's inequality to theoretically prove that when mean travel times are increasing between two time periods, the standard deviation of travel times increases as well during the same time. Other papers such as Kouwenhoven, Schoemakers, Grol, and Kroes (2005) show empirically that a higher mean travel time usually implies a higher level of variability. For the most part, Eliasson (2006) derives similar results. But he also shows that if congestion is very severe, variability can become a decreasing function of travel time.

Also in the current paper we find a non-linear relationship between travel time variability and mean travel time. We investigate whether the non-linearity can be attributed to different 'traffic regimes'. The regimes can be defined according to their location in the speed-flow relationship of traffic. This relation is part of the fundamental diagram of traffic which was first found by Greenshields (1935), and later extended and refined in multiple ways (e.g. Haight, 1963; Hall, 2002; Hall, Hurdle, and Banks, 1992; Kerner, 2004). In this paper we distinguish between the free-flow, congested and hyper-congested regime. The free-flow regime refers to traffic conditions characterized by travel speed close to free-flow speed, and flow rates considerably lower than the maximum capacity of a road segment. The congested

regime describes conditions where travel times are significantly higher than free-flow travel times, and flow is close to maximum capacity. Finally, hyper-congestion occurs if the density that is consistent with maximum flow (capacity) is exceeded and queuing takes place, leading to a below-maximum flow and considerable delays. In terms of the speed-flow curve (plotting traffic flow on the horizontal and speed on the vertical axis), the free-flow and congested regimes are located at the upper (downward bending) segment of the curve, whereas the hyper-congested regime corresponds with the lower (upward bending) section of the curve. Only few papers have looked explicitly at the influence of flow on travel time variability. Tu, Van Lint, and Van Zuylen (2007b) investigate the correlation between (in)flow variability and travel time variability. They find that variability is hardly related to the variability of flow in the free-flow and hyper-congested regime, whereas it is positively correlated with flow variability in the congested regime.[1]. In the current paper, in contrast, we focus on the impact of the relative shares of the traffic regimes on variability.

The current paper also introduces two measures of travel time variability. They differ in the assumption on extent of information available to drivers. The first one is based on the usual assumption that for a given road link, and a given time of the day, expected travel times are equal over the entire year. We will call this "rough information" (RI). The second measure assumes that drivers adjust expectations on travel times according to (public) information on weather conditions, weekday, season and demand patterns ("fine information" (FI)). It has, for example, been shown in earlier research that adverse weather effects do not only cause travel times to be longer, but also increase variability (Tu, Van Lint, and Van Zuylen, 2007a).

This paper contributes to the existing literature on predicting travel time variability by explaining the linkage of travel time variability and mean delay for different traffic regimes. The results give a clear indication that the relative shares of the traffic regimes have an impact on the relation between mean delays and standard deviation. Unlike earlier literature, this paper provides two measures of variability, and has a strong focus on the applicability of the results in CBA.

The paper is organized as follows. Section 2 provides an overview of the travel time data and the variables used in the analysis. Section 3 contains the empirical analyses. A non-traffic-regime based model (NTR) as well as a traffic-regime based model (NT) are derived, and results are analyzed. Section 4 proceeds by discussing the implications of the results for cost-benefit analysis. Finally, section 5 provides some concluding remarks, as well as suggestions for further research.

---

[1] Tu et al. (2007b) define three traffic regimes: fluent traffic, transition traffic and capacity traffic. The definitions of correspond closely to the ones used in this paper (free-flow, congested and hyper-congested traffic).

2   Data and Variables

The goal of the paper is to identify factors that determine travel time variability and to estimate their influence on variability. The focus is here on variables that lend themselves for inclusion in CBA. That is, variables should be likely to be readily available to cost-benefit-analysts (for instance as output of traffic models). Besides this, the level of disaggregation in variables should not be too high. CBA are usually used as an instrument to assess the costs and benefits of relatively large-scaled projects, and our study aims at providing insights into how changes in travel time variability can be addressed in such assessments. Therefore, cost calculations at very disaggregate levels are usually neither useful nor feasible. This, for instance, means that we focus on general indicators of road geometry (such as the number of lanes) rather than more specific ones (as for instance road curvature).
Given our concern with CBA, data will be used that as closely as possible reflect the knowledge that drivers may have. Travel time variability is a concept that is very sensitive towards the information and knowledge of drivers. An underestimation of the information and knowledge available to drivers is equivalent to an overestimation of the variability drivers truly experience, and vice versa.

2.1   Travel Time Data

For the econometric analysis we use travel time distributions of 145 (one-directional) highway links in the Netherlands. Since travel time data are not available for door-to-door trips, we focus on the highway network, which is very dense in the Netherlands. For this reason, almost 50% of all driving takes place on highways.[2] Figure 1 shows a map of the Netherlands, indicating all road stretches included in the analysis. Most of them are concentrated in the West of the country, which is one of the most densely populated areas in Europe, known as Randstad. It includes major cities such as Amsterdam, Rotterdam, The Hague, and Utrecht.

Links have been defined in a way such that each of them starts and ends after a highway intersection. The link length varies between 2.2 km and 37.1 km, with an average of 13.2 km. To define road stretches in this way is partly motivated by the idea that for drivers these intersection points are natural points of reference. Frequently they are bottlenecks and in many cases they are also distinct landmarks. Not choosing all road links to be of equal length also has the additional advantage that it allows to investigate the effect of the length on variability, and renders the research applicable to a larger variety of roads. On all road links that enter the analysis, the density of loop detectors is fairly high. The speed measurements collected from all loop detectors situated at a specific link are aggregated

---

[2]This number refers to non-freight transport. The percentage is based on the number of km driven on Dutch highways in 2006 (around 50 billion km, assuming a 20% rate of freight transport) divided by the number of overall km driven (around 95 billion km).(Ministerie van Verkeer een Waterstaat (Dutch Ministry of Transport), 2006, 2008)

Figure 1: Network of highway links included in the analysis
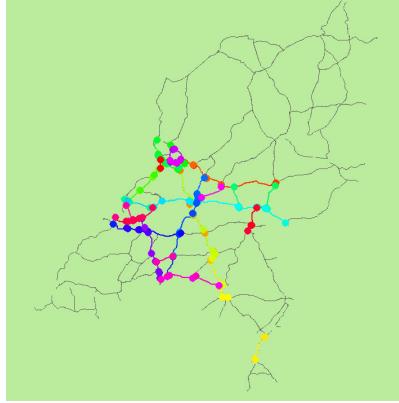
into travel times on the entire link using the piecewise-linear-speed-based (PLSB) trajectory algorithm developed by Van Lint and Zijpp (2003). For each 15-minutes time interval and each link, an average travel time is generated, implying that differences in travel times between individual vehicles that depart within the same interval (vehicle-to-vehicle variability) are not taken into account.

The interval length is chosen to be 15 minutes. This reflects the assumption that travelers are able to distinguish average travel conditions between two adjacent quarters of an hour. The dataset comprises travel time data of all working days in 2008 (255 days). Weekends and public holidays have been excluded as demand patterns are usually quite different on such days. As a vast majority of drivers is aware of these differences, the inclusion of these non-working days would thus cause an overestimation of travel time variability. In addition, we only consider the time span between 6:00 and 20:15 (57 quarters). It is likely that observed travel time variability during night hours is not caused by delays due to congestion but rather by day-to-day differences in the composition of the vehicles that are observed. For instance, an above-average number of freight vehicles might lead to travel time observations that are higher-than-average. This type of variability, however, does not result in variability-induced costs to the drivers and is therefore not useful to include in cost-benefit analyses. Thus, for each link, 57 travel time distributions are defined, each comprising the travel time observations on the 255 working-days of 2008 for a specific quarter of an hour.

2.2 Travel Time Variability

In this paper the standard deviation of travel times is used to represent variability. For an overview of alternative indicators of dispersion we refer to Van Lint, Van Zuylen, and Tu (2008). In the transport economics literature, the unreliability of travel times is often expressed in terms of expected schedule delays (with respect to a preferred arrival time). We cannot use this approach here, as we do not observe the preferred arrival times of drivers

in our sample. They can only be derived implicitly by assuming certain (distributions of) the values attached to travel times and schedule delays. Recent literature (Fosgerau and Karlström, 2010; Li et al., 2009) has shown that under certain conditions, the knowledge of the preferred arrival times is not necessary. In case of a constant (standardized) travel time distribution, the costs related to arriving early or late can be expressed in terms of the standard deviation. In these cases a direct correspondence exists between the approach used in this paper and the approach that involves scheduling costs.

We propose to distinguish between two different measures of variability, both involving the standard deviation of travel times (SD). The distinction stems from the fact that travel time variability can either be based on what we will call "rough information" (RI), versus "fine information" (FI). In the RI case we assume that expected travel times only vary between links and quarters of an hour. In contrast, the FI case assumes that travelers adjust their expectations on travel times also according to weather conditions, season, weekdays and network-wide delays.

It is clear that the two approaches differ in the underlying assumptions on what information an individual uses when constructing an expectation for a given time of the day on a specific day. Introspection suggests that FI is more plausible. Still, whether predictions of variability based on RI or FI are used in CBA depends on the context within which the value of reliability has been derived. Predictions and valuations of travel time variability should be consistent. A value derived in an FI context and applied to an RI prediction of variability, would most likely tend to overestimate the corresponding costs.

The regarding formulas for travel time variability based on RI and FI are shown below:

$$SD_{RI,q,l} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (tt_{i,q,l} - \mu_{q,l})^2} \tag{1}$$

$$SD_{FI,q,l} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (tt_{i,q,l} - \mu_{i,q,l})^2}, \tag{2}$$

where the index q indicates a specific quarter of an hour, l a specific road link, and i stands for a specific day of the year. N is the number of days (equal to 255 here), tt the travel time, and μ the expected travel time. In the case of RI, μ is constant for a certain quarter and link. For FI, μ also depends on day-specific characteristics.

Day-specific expected travel times are determined using regression analysis. Separate regressions are run for each link and each quarter, as it is likely that the effects of the day-specific factors differ across road links and times of the day. The dependent variable in these regression are the travel times ($tt_{i,q,l}$) during workdays of 2008. The explanatory variables used in the regressions are listed in Table 1. First, dummies for weekdays are used. Especially frequent drivers are aware of possible differences in traffic conditions on different days of the week. To capture seasonal effects, the year is divided into three parts; the holiday months July and August, the winter months December and January, and the rest of the year.

Weather-related variables that enter the regression analysis are aggregated over the day. We test the significance of the maximum temperature being below 5°, and above 25°, respectively. Also the effects of a rain duration longer than 5 hours, and of a rain amount exceeding 10 millimeters are tested. At the geographical level, each road-link is connected to the closest of four weather stations[3]. Not letting weather conditions depend on the time of the day and choosing a limited number of weather stations, is consistent with drivers usually not having precise information on weather conditions.

Network-wide delays are an indicator of average congestion on all road links during a specific day. They are based on the sum of delays occurring between 6:00 and 20:15 on all road links during a specific day. The aggregate delays are weighted by the relevance of the link in the network, expressed by annual flow levels (again between 6:00 and 20:15). As a last step, an indicator is formed adjusting the scale of the network-wide delays such that the mean across all days is equal to 1. Network-wide delays are included to reflect more predictable daily demand shocks, causing travel time variations other through the supply factors already identified. Examples are major incidents affecting substantial parts of the network, or extraordinary demand conditions such as on days that are close to national holidays.

Table 1: FI: explanatory variables

Explantory Variables		
Category	Indicator	Type of Variable
Weekdays	Tuesday	dummy
	Wednesday	dummy
	Thursday	dummy
	Friday	dummy
Season	July/August	dummy
	December/January	dummy
Weather	maximum temp. $< 5°$	dummy
	maximum temp. $> 25°$	dummy
	duration rainfall > 5 h	dummy
	amount rainfall > 10mm	dummy
Network	Network-wide Delays	continuous

The regressions carried out for each link and each quarter, using the explanatory variables of Table 1, yield an average R-squared of 0.16 (median: 0.12). Table 6 in the appendix of this paper shows that the R-squared varies considerably across regressions. So, the 10th percentile value corresponds to an R-squared of only 0.03, whereas the 90th percentile value equals 0.34.

For the coefficient medians as well as the 10th and 90th percentiles we again refer to Table 6. This table also shows in how many percent of regressions a certain variable has shown to be

---

[3]Data source: Royal Dutch Meteorological Institute (KNMI). The weather stations are located in Schiphol (Airport Amsterdam), Rotterdam, De Bilt and Gilze-Rijen.

significant at the 1%, 5% and 10% level, respectively. While all coefficients are significant at the 10% level in at least 10 % of all cases, this percentage is considerably higher in the case of the constant term (97%), weekday dummies (42–51%) and the network-wide delays (72%).

Also coefficient values are quite heterogeneous across regressions. Only in the case of the constant term and the network-wide delays, coefficients tend to have the same sign (positive in these two cases). For the other coefficients, medians are fairly close to 0, and coefficients are positive or negative in a similar number of regressions. This results implies counterintuitive results such as decreasing travel times during adverse weather conditions in some regressions. It can be shown, that this is, to a major extent, the consequence of including the variable 'network-wide delays', as this variable partly also captures weather-related effects. Based on the coefficients derived in the regressions, expected travel times can be calculated for each day, using the fitted values. These are an indication of the travel times expectations of drivers who have some knowledge and information on weekday, season, weather and demand-related factors, and their effects on travel times. Based on these day-specific expected travel times (FI), the standard deviation is derived according to Equation (2).

2.3 Mean Delay

It has been shown in previous research that travel time variability is strongly correlated to mean travel times (e.g. Fosgerau, 2010). To include travel times in our analysis, the values first need to be made comparable between links that differ in lengths and free-flow times. Otherwise, a travel time of 30 minutes might indicate free-flow conditions on one road link, and hyper-congested conditions on another link. We found that good results can be obtained when travel times are simply expressed in terms of delays.[4] In order to calculate delays, free-flow times must be determined. We assume that free-flow times are stable over the day.[5] Free-flow times are derived using the numerical optimization method developed by Aerde and Rakha (1995).[6] A speed-flow curve is fitted to traffic observations minimizing the (squared) errors in respect to speed, flow as well as density. The fitting procedure then yields estimates of the free-flow speed (in addition to estimates of the maximum capacity and the speed at the maximum capacity). For the dataset used in the current paper, both travel time and flow data are available. A fitting procedure is run for each link, using all travel time and flow observations available between 6:00 and 20:15 for all working-days

---

[4]Alternative procedures to standardize travel times have been investigated as well, however, produced less powerful results. One alternative is to scale all road stretches to a specific length (e.g. Kouwenhoven et al., 2005). A second alternative is to use average speeds instead of average travel times. And third, a ratio of travel times divided by free-flow time can be used to make travel times on different links comparable (Eliasson, 2006).

[5]This assumption is reasonable for most road links used in the analysis. Only few of them are subject to variable speed-limits, which clearly affect free-flow times.

[6]The according software (SPD-CAL) can be found under: `http://filebox.vt.edu/users/hrakha/Software.htm#SpdCal`

in 2008. As a result, a constant free-flow speed is obtained for each link. The difference between mean travel times and free-flow times gives mean delays. The idea that a 5-minute delay on a road link of 5 km is different from a 5-minute delay on a road link of 30 km is controlled for by including the length of the road link (and various interaction terms with link length) as explanatory variable in the regression analysis.

It should be noted at this point that mean delay is calculated in the same way for both information regimes, FI and RI. This implies that free-flow time is regarded as independent from seasons, weekdays, weather and demand conditions. This assumption is valid in most cases, as the minimum feasible travel time is only likely to become structurally longer under extreme conditions in these factors (e.g. persistent snowfall).

2.4 Link-specific variables

A number of link-specific variables are accounted for. First, the number of lanes and the length of a link are included in the analysis in order to test whether they have an effect on travel time variability. These two variables have been selected as they can be considered exogenous with respect to travel time variability. The number of lanes is to a major extent determined by demand, and the length of the link has been determined by the location of intersection points. Variables such as speed-limits or traffic management measures, in contrast, exhibit a higher degree of endogeneity as they are frequently used to improve travel time variability. Therefore, they do not allow for an unbiased identification of causal relationship regarding variability, which is essential for CBA. Other variables, such as the percentage of freight transport or the number of ramps per kilometer, were tested but did not prove to be significant.

3 Empirical Analysis of Travel Time Variability

The empirical analysis consists of two parts. In the first part, models are estimated that explain variability by mean delay as well as certain road characteristics, but not by traffic regime; thus only using variables that are likely to be available to cost-benefit analysts. These models are named 'non-traffic-regime based' (NTR). From these regressions, we find that the relation between mean delay and variability is positive but non-linear. In the second part, we show that these non-linearities can be largely attributed to differences in the relative shares of the traffic regimes (free-flow, congested and hyper-congested). The corresponding models are summarized under the heading of 'traffic-regime based' (TR).

3.1   Non-traffic-regime-based Model (NTR)

In this section of the paper the relation between mean delay and the standard deviation of travel times is investigated, not taking into account the relative shares of the traffic regimes. Figure 2 plots the mean delay and standard deviation for all working days, for each quarter of an hour for each link. Therefore, the graph contains 8265 observations (145 links times 57 quarters). The graphs shows that the relation between mean delay and the standard deviation is positive, and slightly concave. A higher mean delay is on average associated with higher variability. This intuitively makes sense. Deviations from mean travel times due to fluctuations in road supply and demand are likely to occur more often during times at which delays take place regularly (like morning and evening peaks). In addition, they tend to have a higher impact. Under free-flow conditions, moderate fluctuations in supply conditions have a more limited effect on travel times, since there is still unused capacity.

Figure 2 also demonstrates the difference between RI and FI based variability. In the rough information (RI) case, variabilities are considerably higher than in the fine information (FI) case. This does not come as a surprise, since the average R-squared of the auxiliary regressions (with weather, season, weekday and demand conditions as explanatory variables) was 0.16. As a consequence, that part of travel time variability that is due to the indicators we used for weather, season, weekday and demand conditions has been removed in the FI case, resulting in lower variability.

Furthermore, it is evident from the graphs that the relation between mean delay and travel time variability is not linear. We will examine this further in the regressions below.

Figure 2: Relating Mean Delay and Standard Deviation
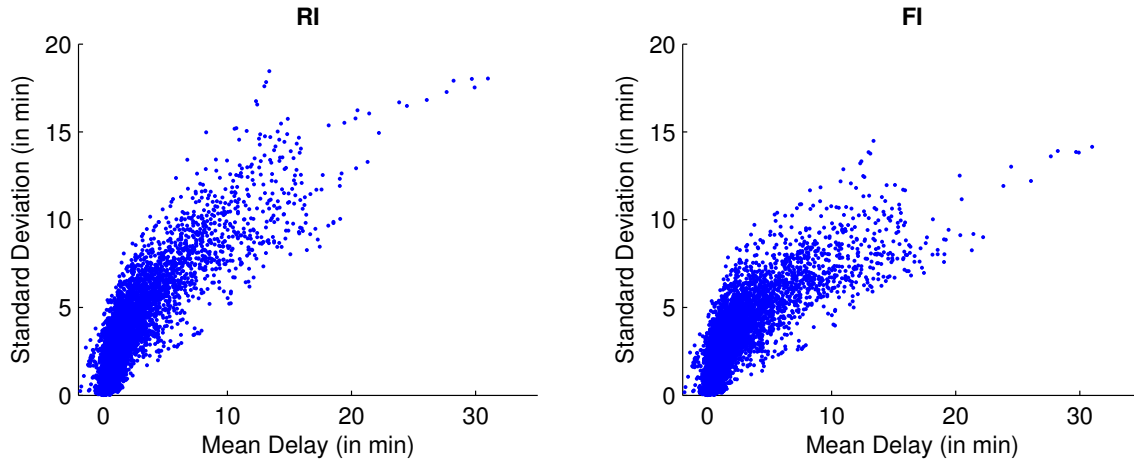
The data points displayed in Figure 2 are also the ones used in the regression analyses. Travel time variability is the dependent variable, and mean delay (together with other variables) shows up at the right hand side of the equation. The panel structure of this dataset,

with the road links as cross-sectional dimension and the quarters as temporal dimension, is taken into account in the form of random effect models. Using a panel structure in the estimations leads to a better fit of the models. We do not use fixed effects models, since they do not allow for the estimation of time-invariant variables. They would absorb all road-specific, time invariant characteristics (such as the number of lanes) in the link-specific error terms. Random effects models in contrast allow for the estimation of such variables as they assume that the link specific error terms result from random, independent drawings of a larger population of road links (Heij, Boer, Franses, Kloek, and Van Dijk, 2004). This assumption is quite realistic in our case, since indeed specific road links have been chosen out of a more extensive highway network. It was verified that the estimates of the (time-variant) variables in the fixed effects model, and the random effects model are very close to each other.

In the first two regressions in Table 2, Model 1, we assume a linear relationship between standard deviation and mean delay. The RI-based model yields an R-squared of 0.75, and the FI-based model an R-squared 0.67. As expected, we find a significantly lower coefficient for mean delay when the standard deviation is based on FI compared to RI (0.58 vs. 0.76). The seemingly contradictory outcome that both the root-mean-squared error (RMSE) and the R-squared are lower in the FI case compared to the RI case, can be explained by the generally lower absolute values of variability in the FI case. The fit of the FI case is worse compared to the RI case. This results from using fitted instead of observed travel time data for the calculation of the FI based variability. It can be verified that including more day-specific explanatory variables in the auxiliary regressions (in addition to the ones shown in Table 1 improves the fit of the FI-based models.

The third and fourth regression in Table 2, Model 2, assume a non-linear relationship between standard deviation and mean delay. In addition to mean delay, average speed, the length of the road link (in km), the average number of lanes, and various interaction and exponential terms of these variables are included as explanatory variables. Also free-flow speed and speed-at-capacity are added as regressors. Speed-at-capacity indicates the speed that corresponds to the maximum capacity of a road link on the speed-flow curve. Like free-flow speed, also speed-at-capacity is derived using the methodology developed by Aerde and Rakha (1995). Average traffic flow is not included in the equation as it is double-valued with respect to travel time for a conventional backward-bending speed-flow relation. Flows can be small for two reasons. They might be small as a consequence of low demand for travel, but also as a result of high demand, leading to slowly moving traffic jams.

We find that the fit of Model 2 is considerably better compared to Model 1. The RMSE drops from 1.13 to 0.89 for RI and from 1.05 to 0.82 for FI, respectively. Similar to Model 1, the fit is better in the RI case, despite of the RMSE being higher than in the FI case.

The variables used in Model 2 are partly related to each other. Speed is defined as link length divided by mean travel time (in hours). Mean travel time, in turn, is defined as the sum of delay and travel time under free-flow conditions. Taking into account the interactions between these variables, it can be shown that the derivative of the standard deviation with respect to mean delay decreases in mean delay. The regressions shown in the subsequent section will show that this result can be attributed to differences in the relative shares of

Table 2: Regression Results (Non-traffic-regime-based Model)

	Model 3		Model 4	
VARIABLES	RI	FI	RI	FI
Mean Delay (MD)	0.764***	0.578***	1.319***	1.191***
MD^2			-0.040***	-0.048***
MD^3			$6.51 * 10^{-4}$***	$9.47 * 10^{-4}$***
Mean Speed (MS)			0.187***	0.183***
MS^2			$-1.28 * 10^{-3}$***	$-1.21 * 10^{-3}$***
Length			0.152***	0.140***
Length^2			$-3.20 * 10^{-3}$***	$-2.84 * 10^{-3}$***
MD*Length			$-1.47 * 10^{-3}$*	$-4.12 * 10^{-3}$***
Lanes			0.172**	0.147*
MD*Lanes			-0.053***	-0.026***
Free-Flow Speed			0.021***	0.013*
Speed-at-Capacity			0.018**	0.015***
Constant	1.451***	1.455***	-10.260***	-9.312***
R-squared	0.751	0.667	0.856	0.810
RMSE	1.126	1.052	0.889	0.824

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

traffic regimes.

We also find that *ceteris paribus*, a shorter link is associated with lower variability. This result is reasonable, as variations in demand and road capacity tend to cause smaller deviations from average travel times on shorter roads compared to longer roads. Also the number of incidents can be assumed to be higher on longer roads. Nevertheless, variability increases less than proportionally to link length. This can be explained by 'averaging-out' effects. On longer road links, delays on one part of the link can be compensated by free-flow conditions on another part of the link. For instance, after an accident has reduced road capacity, queues will form upstream of the incident location, whereas downstream free-flow conditions are likely.

Also the marginal effect of the number of lanes on variability is a function of mean delay. For smaller delays, variability is positively correlated with the number of lanes. This may be caused by the fact that a higher number of lanes induces more lane changes. But especially under high speeds, a relatively high accident risk is related to lane changes leading to increased variability. For longer delays, however, variability is negatively correlated with the number of lanes. It is likely that under more slowly moving traffic conditions, lane changes are associated with a smaller accident risk compared to fast moving traffic. Indeed another effect becomes more important. If an incident happens, a higher number of lanes gives more possibilities for traffic to still pass, decreasing variability.

Both free-flow speed as well as speed-at-capacity are correlated positively with variability.[7]

---

[7]Note that free-flow speed does not only enter the equation directly but also indirectly, via the calculation of mean speed. A higher free-flow speed is associated with higher mean speed. Mean speed in turn is correlated positively with variability (Tabel 2). So, both the direct and the indirect relation between free-flow

It can be shown that a positive relation exists between free-flow speed and speed-at-capacity. Roads with both low free-flow-speed and low speed-at-capacity tend to be roads that are easily congestible. Already at relatively high speeds, congestion and hyper-congestion, respectively, appear, implying also higher levels of variability. Reasons for lower values of free-flow-speed and speed-at-capacity are manifold. Road geometry (changes in the number of lanes, curves with small radius, etc.) might be one reason. Traffic management measures such as speed-limits or the use of the shoulder as driving lane might be another one. Also spill-back effects from downstream links are possible to cause free-flow speed and speed-at capacity to differ between links.

3.2 Traffic-Regime based Model (TR)

From the outcomes of Model 2, we found that the derivative of the standard deviation with respect to mean delay decreases in mean delay. Figure 3 illustrates that this pattern is related closely to traffic regimes. We distinguish three traffic regimes: free-flow, congested and hyper-congested. Free-flow conditions are characterized by a speed that is higher than a fraction 0.9 of the speed associated with free-flow travel time. Congested conditions are defined for lower speeds, but higher than speed-at-capacity. Hyper-congested conditions prevail if the speed drops below speed-at-capacity. Therefore, flow congestion is present in the free-flow and the congested regime, and (downstream) bottleneck congestion in the hyper-congested regime[8]. In Figure 3, each data point shown in Figure 2 is attributed to the dominant traffic regime in terms of relative shares. This means that if for a certain road-stretch, time-of-day combination, travel times have most often been observed in the free-flow regime, the corresponding data point is shown in the upper left of Figure 3.

We find that the slope is steeper in the free-flow regime compared to the congested regime, and that the slope of the congested regime is steeper than the slope in the hyper-congested regime. This gives a possible explanation for the non-linear distribution of data points found in Figure 2. Since the average number of observations with hyper-congestion as dominant regime is increasing in mean delay, the slope becomes less steep at higher levels of mean delay. For the single graphs shown in Figure 3, slopes are much closer to linear compared to that what is suggested in Figure 2.

Again we use regression analysis to test these notions. As for the NTR models, random effects are used. In the first two estimations in Table 3, Model (3), mean delay is interacted with the three regimes. In Model 4, the additional regressors from Model 2 are added.

Model 3 includes MD (mean delay) interacted with the shares of the three regimes among the 255 observations for each link-daytime combination. Comparing the fit of Model 3 to the NTR models, we find that it does not only have a better fit compared to Model 1 but also compared to Model 2. Compared to Model 2, the RMSE drops from 0.89 (0.82) to

---

speed and variability are positive.

[8]Although results do not change qualitatively, the distinction into three rather than two regimes (congested vs. hyper-congested) leads to a substantial improvement in the fit of the models.

Figure 3: Traffic regimes
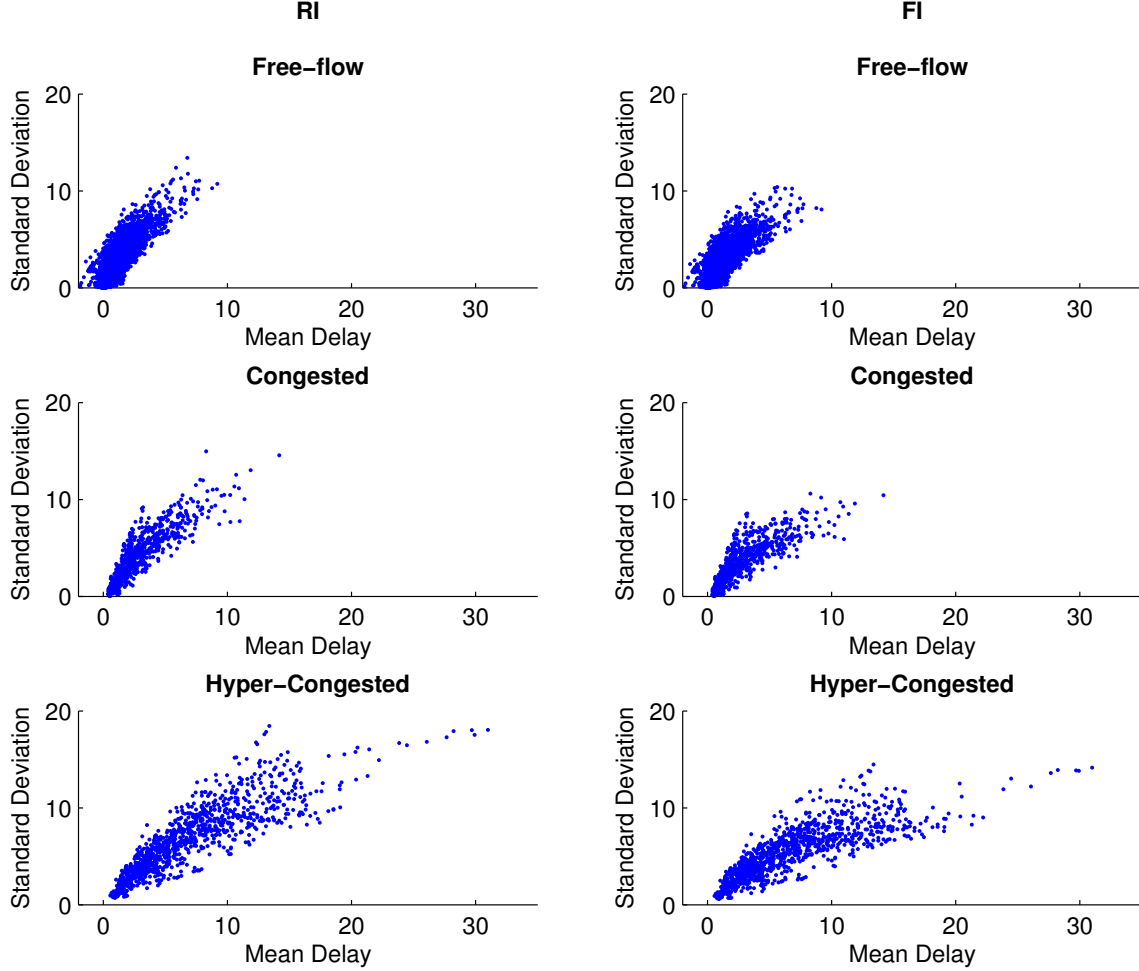
0.78 (0.78) for the RI (FI) case. This is an indication of the power of the traffic regimes to actually explain and predict travel time variability. Consistent with Figure 3 we find that the lower the regime is ranked in terms of speed, the lower the derivative of the standard deviation with respect to mean delay.

A possible explanation of this finding is that for quarters in which on most days no congestion occurs, mean delays are strongly associated with (non-recurrent) fluctuations in demand and supply. In contrast, during time intervals in which hyper-congested conditions are dominant, the observed mean delays are likely to be strongly associated with recurrent demand patterns. Comparing these two sources of delays, fluctuations in demand and supply are associated with relatively more variability compared to recurrent congestion patterns. Clearly, fluctuations in demand and supply also occur under circumstances where hyper-

congestion is the usual traffic condition. However, since we define the road links over several kilometers, leading from one highway intersection to another, and allow for highway entries and exits on the links, the links can include one or more bottlenecks. Traffic conditions are therefore likely to be heterogeneous across a link. Upstream of the bottleneck(s), hyper-congested traffic conditions can be found, whereas downstream of the bottleneck, the congested and free-flow regime prevail. As a consequence, delays can be expected to average out over a link. If an incident causes one of the upstream bottlenecks to have unusually long queues, the smaller outflow may imply shorter queues at downstream bottlenecks. Thus, demand and supply fluctuations are suggested to have a relatively smaller impact on travel times and travel time variability on a, normally, hyper-congested link than on a link that is usually subject to free-flow conditions. This gives a possible explanation for the comparatively lower coefficient of mean delay interacted with the share of hyper-congested observations.

In Model 4, the additional explanatory variables from Model 2 are added to Model 3. The coefficients in Table 3 give an indication that the traffic regimes capture the non-linearities in the relationship of the standard deviation and mean delay. Adding the explanatory variables of Model 2 to the regimes interacted with delay, leads to only minor improvements in terms of RMSE (compared to Model 3). For the RI-based models the RMSE decreases by 0.02, and for the FI-based models by 0.01. The relative improvement from adding additional regressors were considerably higher for the NTR models, where the relative declines were equal to 0.23 for both the RI and the FI case.

The coefficients of the traffic regimes interacted with mean delay are quite similar in Models 3 and 4, especially for the free-flow and the congested regime. The differences between the coefficients for mean delay interacted with the share of hyper-congested observations is a result of including the second and the third power of mean delay as regressors, as these are strongly correlated to the share of hyper-congested observations.

Compared to Model 2, the additional regressors in Model 4 show, on average, a lower significance. Their signs, however, are in most cases equal to the ones observed in Model 2. Changes in signs are an indication that the relative shares of the traffic regimes capture the impact of the regarding variable on variability.

4   Implications for CBA

From the results in the previous section of the paper, we find that the TR models yield the best predictions of variability. Compared to the NTR models, they have a lower RMSE and higher explanatory power, and are, therefore, also the preferred model to use in CBA. However, using NTR models for the predicting variability after specific projects have been implemented, requires the cost-benefit analyst to know how the relative shares of the traffic regimes will change due to these projects. For instance, in the case of changes in relative capacities of upstream and downstream links, one might use the size of these changes to infer changes in the relative shares of the traffic regimes. If downstream capacity becomes

Table 3: Regression Results (Traffic-regime-based Model)

| VARIABLES	Model 3		Model 4	
	RI	FI	RI	FI
MD*Share(free-flow)	2.291***	1.983***	2.180***	1.850***
MD*Share(congested)	1.365***	0.998***	1.343***	1.071***
MD*Share(hyper-congested)	0.414***	0.287***	0.190***	0.384***
MD^2			0.034***	$5.46 * 10^{-3*}$
MD^3			$-6.92 * 10^{-4***}$	$-7.44 * 10^{-5}$
Mean Speed (MS)			0.035***	0.073***
MS^2			$-3.40 * 10^{-4***}$	$-5.42 * 10^{-4***}$
Length			0.125***	0.120***
$Length^2$			$-3.14 * 10^{-3***}$	$-2.85 * 10^{-3***}$
MD*Length			$-0.013***$	$-0.011***$
Lanes			0.038	0.047
MD*Lanes			0.010	0.013*
Free-Flow Speed			$-8.33 * 10^{-3}$	$-5.76 * 10^{-3}$
Speed-at-Capacity			0.025***	0.018***
Constant	0.480***	0.589***	$-1.624**$	$-3.164***$
R-squared	0.864	0.808	0.896	0.846
RMSE	0.781	0.775	0.758	0.749

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

relatively smaller, hyper-congestion is more likely to occur. In contrast, if upstream capacity becomes relatively smaller, free-flow conditions are more probable.

However, in most cases the changes in relative shares of the traffic regimes are difficult to predict. In these cases Model 2 is the preferred alternative. It shows a significantly better fit than Model 1, which assumes a linear relation between mean delay and variability. And it includes variables which are usually available to cost-benefit analysts. Moreover, Figure 4 verifies that there is no structural bias in the estimation of Model 2. For most out of the 8265 observations, the difference between fitted and actual values of variability is close to 0.

4.1　An Example based on Model 2

In this section a numerical example based on Model 2 and its implications for CBA are discussed. Based on the coefficients derived in Table 2, predictions of variability are presented as a function of mean delay as well as road link length for both RI and FI. Road link length is added since it shows to have considerable impact on variability. For the other variables, values close to their sample means are used ((average) number of lanes=2.5; free-flow speed=105 km/h; speed-at-capacity=80 km/h). Figure 5 shows 3D graphs of the variability predictions for the RI- and FI-based models. Tables 4 and 5 (in the appendix) summarize the underlying numerical data of the predictions, as well as of the first and the second derivative of variability with respect to mean delay.

Figure 4: Difference of fitted and actual values of variability
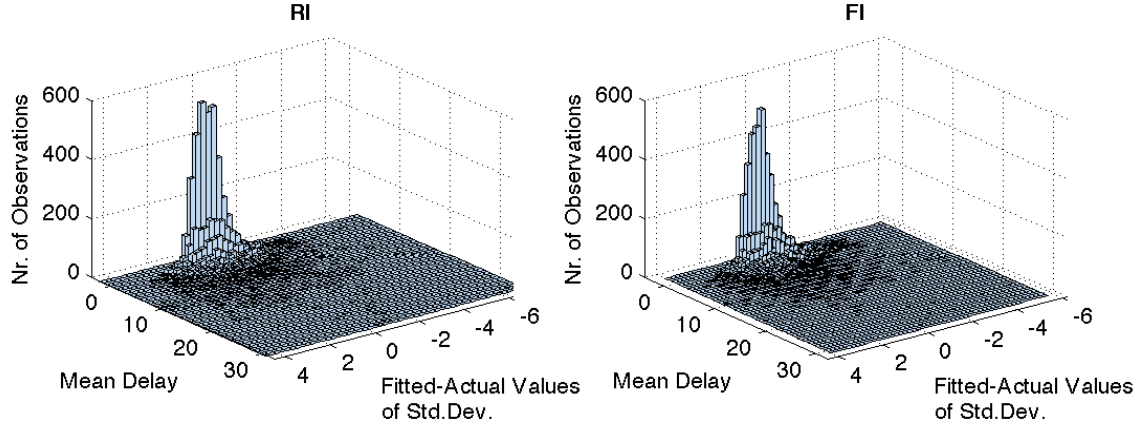
It is confirmed that the first derivative is a decreasing function of mean delay.[9] Also it can be clearly shown that variability is larger for longer road links (however, less than proportionally). Comparing the RI and the FI case, we can observed that the gap between the regarding variability predictions increases in mean delay, both in absolute and relative terms.

Figure 5: Model 2
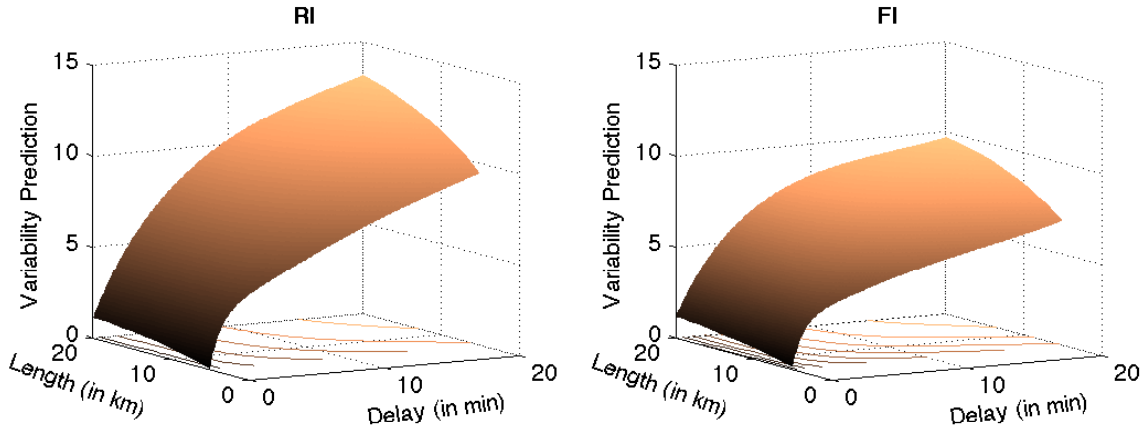
As for CBA the absolute levels of travel time variability probably matter less than changes in travel time variability, the non-linear shape of the relationship between mean delay and standard deviation has an intriguing consequence: If delay is high and it increases

---

[9]The first derivative increases again slightly for longer mean delays. This can be considered a result from very few observations in these regimes. Only 0.7% of the mean delays in the sample exceed 15 minutes.

further, this leads to a smaller absolute increase in travel time variability than if the initial delay had been smaller. Therefore the finding that the relation between mean delay and standard deviation is close to linear for somewhat longer delays, is especially interesting in respect to CBA. A linear relationship between the two variables is straightforward to use in CBA, as the rate of change in variability with respect to mean delay does not depend on mean delay. Tables 4 and 5 show that the second derivative becomes close to 0 at delays of around 4 to 8 minutes (depending on the length of the road link). Therefore, for longer delays it is valid to use a constant ratio that directly connects costs of variability with costs of delays.

An illustrative calculation can show the simplicity of this approach. In doing so, we assume a reliability ratio (=value of reliability divided by value of time) of 0.8, which is consistent with what has been found in previous research (e.g. Hamer, Jong, Kroes, and Warffemius, 2005). This number can next be multiplied by the first derivatives given in Tables 4 and 5, and associated with longer delays. For instance, for a delay of 8 minutes on a link of 10 km, the increase in the costs of reliability if costs of travel time change by 1 is equal to $0.8 * 0.41 = 0.32$ under the assumption of RI, and $0.8 * 0.24 = 0.19$ under the assumption of FI. These numbers are quite close to Dutch cost-benefit guideline of the CPB (Netherlands Bureau for Applied Policy Analysis) which assumes an according value of 0.25. (Besseling, Groot, and Verrips, 2004).

4.2   The Use of FI- and RI-based variability

It remains to discuss whether in CBA the extent of unreliability shall be measured based on RI or FI based variability. Ultimately, CBA is focused on calculating the costs related to variability. For this purpose it requires both predictions of the future development of variability and monetary values that drivers attach to changes in variability. These two factors should of course correspond to each other as closely as possible in order to obtain realistic results. The monetary values attached to variability (or alternatively, schedule delays) are either derived from revealed preference (RP) or stated preference (SP) studies. RP valuations are derived from actual, observed choices of travelers in respect to departure time, route and mode. SP studies, in contrast, are based on hypothetical choice experiments, in which participants are asked to indicate their preferred alternative out of a certain choice set. In terms of variability, these choices usually state two or more possible travel times, which occur with specific probabilities. These differences in the underlying data source also imply differences with respect to the information that is assumed available to drivers.

Drivers participating in RP studies are likely to make their travel choices taking into account some information on day-specific factors, suggesting FI might be more appropriate. However, in order to measure the trade-offs that drivers face, variability on the roads frequented by these drivers needs to be known. In most cases the variability is measured on basis of observed traffic variability, without taking into account day-specific factors. The same idea is reflected in the RI based value measure, which also includes variabilities that drivers are

actually aware of. This makes it more adequate to use RI rather than FI in such cases, even though real behavior is based on FI.

SP studies usually state that travel times in the choice experiments are purely random. For instance, Tseng, Verhoef, De Jong, Kouwenhoven, and Hoorn (2009) compare travel time outcomes to the outcome of rolling a dice. If respondents follow this interpretation of pure randomness, it is more appropriate to measure variability in terms of FI rather than RI, as the FI-based variability widely eliminates variability due to factors usually known to drivers. Using the RI-based variability instead, results in an overestimation of the costs associated with variability. It is likely that in reality CBA tend to overestimate the costs of reliability due this mismatch between the underlying assumptions of the valuation and the volume of variability. However, the extent of overestimation might be smaller than what is expected at first sight. Namely, respondents might might interpret the travel time distributions presented in SP studies as some distribution over all relevant (working) days, so that in reality they would have more information than just this distribution. This results in less uncertainty and, therefore, a lower estimated value of reliability (e.g. Euros per unit of standard deviation). If this effect is large enough, the costs of variability might be represented more truthfully by an RI rather than an FI-based measure of variability.

If and to which extent this effect exists, needs to be judged on a case-by-case basis. It is likely that the information that respondents obtain during the SP experiment on the underlying assumptions of the travel time distribution plays a major role in this respect. The quality of the SP studies and their usefulness for CBA largely depends on whether a concise and understandable explanation on the background of travel time variability is given. Finally note that FI is not uniquely defined. in this paper, FI is based on the factors cited in Table 1. These factors have been chosen as a representation of the information that most drivers have. In reality, the extent and quality of information does not only differ widely across drivers, but is also likely to be different for different times of the day or different road links. The definition of FI based variability can be adjusted for these aspects, if required.

5   Conclusions

We showed that travel time variability is well predictable using econometric models. Variability is correlated positively with delays, and the first derivative of travel time variability with respect to mean delay is decreasing in delay. For longer delays the relation between these two variables is close to linear. The non-linear relation between mean delay and variability can be related to the distribution of traffic regimes. A one minute increase in mean delay is associated with a smaller increase in variability if the relative share of hyper-congested traffic observations is high during that time of the day. It is associated with a higher increase in variability, however, if free-flow observations are dominant.

Eliasson (2006) as well as Fosgerau et al. (2008) come to the conclusion that the costs related to travel time variability are around 10-15% of the costs related to changes in mean travel times. For longer delays we find a value of 19% for RI-based variability and 32%

for FI-based variability (for a road link of 10 km). For shorter delays, costs tend to be higher. One possible explanation for the divergence with prior studies is that the analysis in this paper refers to highways, whereas Eliasson (2006) and Fosgerau et al. (2008) analyze travel time variability on urban roads. In urban networks travel time variability is already at a high level for no or very small delays (e.g. due to traffic lights). As a consequence, an increase in mean delay might increase travel time variability relatively less.

Due to the different characteristics of the road links used in the analysis, the results obtained can be used to predict travel times on other Dutch highways and, with some caution, perhaps also on highways in other countries. In this case it must be kept in mind that the properties of Dutch highways put a restriction on the value of some explanatory variables. For instance, in the Netherlands, the highest admissible speed-limit is 120 km/h, whereas the maximum speed-limit is higher or does not even exist in other countries (e.g. Germany). The analysis in this paper is based on travel time data of highways. Therefore, results should not be applied to roads that are considerably different from highways. However, the analysis done in this paper can be conducted in a similar way based on travel time data gathered from other road types.

One should also keep in mind that for the analysis in this paper, road links were defined between two intersection points. This makes the point estimates considerably less applicable to stretches defined in other ways (e.g. for fixed lengths).

The analysis in this paper emphasizes the aggregate relationship between mean delay and travel time variability. However, the effect of traffic management measures on variability might differ considerably from their effect on travel times in terms of size and/or direction. In the analysis of this paper, these effects are not distinguished. Future research may therefore focus on the impact of such traffic management measures on variability, not only because of scientific relevance, but also for practical reasons. CBA often focuses on the evaluation of such measures.

Acknowledgements

A Appendix

A.1 Variability predictions based on Model 2

Tables 4 and 5 provide the underlying data of Figure 5. For three different link lengths (5, 10 and 20 km), the according variability predictions, and the first and second derivative of the standard deviation with respect to mean delay are presented.

Table 4: Variability Predictions (RI) based on Model 2

| Delay | Length of the Road Link | | | | | | | | |
| | 5 km | | | 10 km | | | 20 km | | |
(in minutes)	f	f'	f"	f	f'	f"	f	f'	f"
0.5	1.57	2.25	-2.55	1.72	2.07	-1.00	2.00	1.71	-0.36
1	2.44	1.34	-1.24	2.64	1.65	-0.70	2.82	1.55	-0.31
2	3.37	0.66	-0.34	4.01	1.14	-0.36	4.23	1.28	-0.23
4	4.34	0.41	-0.03	5.77	0.70	-0.13	6.41	0.93	-0.14
8	5.88	0.37	-0.01	7.88	0.41	-0.04	9.26	0.55	-0.07
16	8.39	0.27	-0.01	10.33	0.24	-0.01	12.20	0.25	-0.02

f: Absolute Prediction of the Standard Deviation (SD),
f': First Derivative (dSD/dMD), f": Second Derivative (d^2SD/dMD^2)
Assumptions: Nr. of Lanes=2.5; Free-flow speed=105 km/h, Speed-at-capacity=80 km/h

Table 5: Variability Predictions (FI) based on Model 2

| Delay | Length of the Road Link | | | | | | | | |
| | 5 km | | | 10 km | | | 20 km | | |
(in minutes)	f	f'	f"	f	f'	f"	f	f'	f"
0.5	1.45	1.94	-2.34	1.63	1.82	-0.93	1.93	1.51	-0.35
1	2.19	1.12	-1.13	2.44	1.43	-0.65	2.64	1.34	-0.30
2	2.93	0.50	-0.30	3.60	0.96	-0.34	3.85	1.08	-0.22
4	3.62	0.27	-0.03	5.02	0.53	-0.13	5.63	0.73	-0.14
8	4.60	0.22	-0.01	6.48	0.25	-0.04	7.69	0.35	-0.06
16	6.06	0.17	0.01	7.75	0.12	0.00	9.19	0.11	0.00

f: Absolute Prediction of the Standard Deviation (SD),
f': First Derivative (dSD/dMD), f": Second Derivative (d^2SD/dMD^2)
Assumptions: Nr. of Lanes=2.5; Free-flow speed=105 km/h, Speed-at-capacity=80 km/h

A.2   Results of the Auxiliary Regressions

Table 6: Results Auxiliary Regressions

VARIABLE	Coefficent			Percentage of Coefficients significant at the		
	Median	10th Perc.	90th Perc.	1% level	5% level	10% level
Constant	6.44	2.46	13.99	94.91	96.19	96.71
Tuesday	-0.02	-1.85	1.65	28.83	39.18	46.47
Wednesday	-0.01	-1.81	2.08	31.52	43.63	51.37
Thursday	0.01	-1.43	1.59	24.82	34.69	42.78
Friday	-0.01	-1.41	1.76	23.98	34.53	41.89
July/August	-0.01	-0.96	0.96	8.53	15.06	19.84
December/January	-0.04	-0.81	0.99	5.83	11.92	16.41
maximum temp. $< 5°$	-0.04	-0.89	0.64	3.74	7.94	11.74
maximum temp. $> 25°$	-0.03	-0.84	1.06	4.45	8.49	12.55
duration rainfall > 5 h	0.01	-0.78	0.99	7.33	15.03	20.22
amount rainfall > 10mm	-0.03	-1.20	1.17	4.53	10.21	16.42
Network-wide Delays	0.61	0.04	5.21	54.80	66.73	72.34
R-Squared	0.12	0.03	0.34			

References

Aerde, M. V., and Rakha, H. (1995). Multivariate calibration of single regime speed-flow-density relationships. *Proceedings of the Vehicle Navigation and Information Conference (VNIS)*, 334–341.

Besseling, P., Groot, W., and Verrips, A. (2004). Economische toets op de nota mobiliteit. *CPB Document*, *65*. http://www.cpb.nl/nl/pub/cpbreeksen/document/65/doc65.pdf.

Eliasson, J. (2006). Forecasting travel time variability. *Proceedings of the European Transport Conference 2006, Strasbourg, France*.

Fosgerau, M. (2010). On the relation between the mean and variance of delay in dynamic queues with random capacity and demand. *Journal of Economic Dynamics and Control*, *34* (4), 598 - 603.

Fosgerau, M., Hjorth, K., Brems, C., and Fukuda, D. (2008). Travel time variability. definition and valuation. *DTU Transport*.

Fosgerau, M., and Karlström, A. (2010). The value of reliability. *Transportation Research Part B*, *44* (1), 38–49.

Greenshields, B. (1935). A study of traffic capacity. *Proceedings of the highway research board*, *14*(1), 448–477.

Haight, F. (1963). *Mathematical theories of traffic flow.* Academic Press, New York.

Hall, F. (2002). Traffic Stream characteristics. In N. Gartner, M. C.J., and Rathi (Eds.), *Traffic flow theory: A state-of-the-art report* (chap. 2). Washington, DC: US Department of Transportation, Turner-Fairbank Highway Research Center. Available from `http://www.tfhrc.gov/its/tft/tft.htm`

Hall, F., Hurdle, V., and Banks, J. (1992). Synthesis of recent work on the nature of speed-flow and flow-occupancy (or density) relationships on freeways.

Hamer, R., Jong, G. D., Kroes, E., and Warffemius, P. (2005). *The value of reliability in transport. Provisional values for the Netherlands based on experts opinion.* Leiden, Netherlands: RAND Europe.

Heij, C., Boer, P. de, Franses, P. H., Kloek, T., and Van Dijk, H. K. (2004). *Econometric methods with applications in business and economics.* Oxford University Press, USA.

Hensher, D. (2001). The valuation of commuter travel time savings for car drivers: Evaluating alternative model specifications. *Transportation*, *28*(2), 101–118.

Kerner, B. (2004). *The physics of traffic: empirical freeway pattern features, engineering applications, and theory.* Springer Verlag.

Kouwenhoven, M., Schoemakers, A., Grol, R. V., and Kroes, E. (2005). Development of a tool to assess the reliability of Dutch road networks. *Association for European Transport and Contributors*.

Lam, T., and Small, K. (2001). The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation Review*, *37*(2-3), 231–251.

Li, H., Bliemer, M., and Bovy, P. (2009). Reliability-based dynamic discrete network design with stochastic networks. *Transportation and Traffic Theory 2009: Golden Jubilee*, 651–673.

Ministerie van Verkeer een Waterstaat (Dutch Ministry of Transport). (2006). Mobiliteitsonderzoek Nederland 2006.

Ministerie van Verkeer een Waterstaat (Dutch Ministry of Transport). (2008). Kwartaalmonitor bereikbaarheidsontwikkeling hoofdwegennet. , *1/2008*.

Nagel, K., and Rasmussen, S. (1995). Traffic at the edge of chaos. *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, 222–235.

Nicholson, A., and Du, Z. (1997). Degradable transportation systems: an integrated equilibrium model. *Transportation Research Part B: Methodological*, *31*(3), 209–223.

Small, K. (1982). The scheduling of consumer activities: work trips. *The American Economic Review*, *72*(3), 467–479.

Tseng, Y., Verhoef, E., De Jong, G., Kouwenhoven, M., and Hoorn, T. Van der. (2009). A pilot study into the perception of unreliability of travel times using in-depth interviews. *Journal of Choice Modelling*, *2*(1), 8.

Tu, H., Van Lint, J., and Van Zuylen, H. (2007a). The impact of adverse weather on travel time variability of freeway corridors. *86th Transportation Research Board*.

Tu, H., Van Lint, J., and Van Zuylen, H. (2007b). Impact of Traffic Flow on Travel Time Variability of Freeway Corridors. *Transportation Research Record: Journal of the Transportation Research Board*, *1993*(-1), 59–66.

Van Lint, J., Van Zuylen, H., and Tu, H. (2008). Travel time unreliability on freeways: Why measures based on variance tell only half the story. *Transportation Research Part A: Policy and Practice*, *42*(1), 258–277.

Van Lint, J., and Zijpp, N. Van der. (2003). An improved travel-time estimation algorithm using dual loop detectors.