

Pasidis, Ilias-Nikiforos

Conference Paper

Congestion by accident? Traffic and accidents in England

55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Pasidis, Ilias-Nikiforos (2015) : Congestion by accident? Traffic and accidents in England, 55th Congress of the European Regional Science Association: "World Renaissance: Changing roles for people and places", 25-28 August 2015, Lisbon, Portugal, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/124809>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Congestion by accident? Traffic and accidents in England.

(Extended Abstract)

Ilias Pasidis[§]

Universitat de Barcelona and Institut d'Economia de Barcelona

February 2015

ABSTRACT: The goal of this paper is the estimation of the effect of accidents on traffic congestion and vice versa. In order to do this, I use "big data" of highway traffic and accidents in England for the period 2007-2013. The data exhibit some remarkably stable cyclical pattern of highway traffic which is used as a research setting that enables the identification of the causal effect of accidents on traffic congestion and vice versa. The estimation draws on panel data methods that have previously been used to analyse the behaviour of electricity day-ahead market prices. A positive relationship between traffic congestion and road accidents would yield multiplicative benefits for policies that aim at reducing either of them. In addition, we will explore potential heterogeneous effects around different metropolitan areas and regions. This is one of the few studies that makes use of the increasing volume of big datasets, which are publicly available from governments and local authorities.

Key words: accidents, traffic, congestion, highways, big data, panel, England.
JEL classification: R4

[§]Department of Public Economics, Universitat de Barcelona, Avinguda Diagonal 690, 08034 Barcelona, Spain (e-mail: ipasidis@ub.edu; phone: +34 68 403 7346).

1 Introduction.

Traffic congestion and road accidents are considered the most important sources of external costs related to car travel (Shefer and Rietveld, 1997). Traffic congestion is an omnipresent phenomenon during rush hour in densely-populated regions (see, for example, Arnott and Small, 1994; Downs, 2004). Congestion is an important problem for road transport and a main challenge for transport policy at all levels. Congestion costs Europe about 1% of Gross Domestic Product (GDP) every year (European Commission, 2011) and its mitigation is the main priority for most infrastructure, traffic management and road charging measures.

Congestion typically occurs at times of high travel demand or as a consequence of accidents or other non-recurring incidents that temporarily reduce a road's capacity. Non-recurrent congestion on highways is mostly caused by road accidents and other types of incidents (e.g., object on road, car breakdown) (Adler et al., 2013). This type of congestion constitutes roughly one-quarter of highway congestion (Snelder et. al, 2013). In addition, close to 50,000 people lost their lives and millions were injured as a result of road accidents in Europe in 2002. The total annual costs for society on the basis of the valuation of accidents presented in the COWI (2006) report was estimated at 229 billion per year.

The goal of this paper is to estimate the *causal* effect of accidents on traffic congestion and vice versa. If a positive relationship between the two is identified, policies that aim at reducing either of the two issues will have multiplicative benefits. The COWI (2006) report, which conducted Economic Cost-Benefit analysis for the DG-TREN of the European Commission, ignored such benefits in its assessment of the 21 vehicle safety technologies that the European Commission promoted in order to reduce the number of road fatalities by 50% in 2010¹. The estimation of the causal effects of highway congestion on highway accidents is thus considered an important topic which has partly been neglected by the policy-makers until now.

Nevertheless, estimating the effect of highway accidents on highway congestion is not straightforward. Road accidents typically occur in high congestion times. At the same time, accidents cause traffic congestion. In addition, both congestion and accidents are affected by several other observable and unobservable factors (e.g. road quality and road conditions). Therefore, the identification of a causal relationship between road congestion and road accidents, or between road accidents and road congestion is particularly cumbersome.

While the literature on the effect of traffic congestion on road accidents dates back to the '70s (Vickrey, 1968, 1969; Jones-Lee, 1990; Newbery, 1987, 1988; Vitaliano and Held, 1991; Jansson, 1994; Shefer and Rietveld, 1997; Dickerson et al., 2000; Golob and Recker, 2003; Noland and Quddus, 2005; Quddus et al., 2010), limited attention has been paid on the inverse relationship. In addition, as a large part of this literature is relatively old, new data sources and modern identification techniques have not been employed in this area yet. The existing literature has mentioned some endogeneity concerns, but these issues have not always been addressed adequately.

¹EC objective documented in the White Paper (European Commission, 2001).

This research will estimate the effect of an accident's occurrence on the observed average² speeds using the observed patterns of traffic flows in England's highways in the period 2007-2013³. The combination of these periodic patterns, together with panel data methods that have previously been used to analyse electricity day-ahead market prices (Huisman et al., 2007), provide the framework for such an estimation with a causal interpretation. For the inverse causal relationship, I will use dynamic panel techniques and GMM. Finally, this is one of the few studies that uses a small portion of this increasing volume of big datasets which becomes available from governments and local authorities worldwide. This can be regarded as an important contribution to the economics literature in general since until now, economists have been reluctant to use "big data" in academic research (Varian, 2014).

2 Data

This paper uses very detailed data on highway traffic and accidents for England that are publicly available from `data.gov.uk`. These data have never been used before in an academic paper based on my knowledge. Sometimes, it is the size of such big datasets that is considered an issue but in most of the cases it is the detail of their information that is regarded as superfluous. However, the volume of information in the highway traffic dataset reveals some interesting patterns that allow the identification of the causal effect of highway accidents on traffic speeds and vice versa.

The Highways Agency network journey time and traffic flow data series provide average journey time, speed and traffic flow information⁴ for 15-minute periods since April 2009 to September 2014 on all motorways and 'A' roads managed by the Highways Agency, known as the Strategic Road Network, in England. Journey times⁵ and speeds are estimated using a combination of sources, including Automatic Number Plate Recognition (ANPR) cameras, in-vehicle Global Positioning Systems (GPS) and inductive loops built into the road surface.

The accidents dataset provides detailed information about the circumstances of personal injury road accidents in Great Britain from 2005 onwards, the types of vehicles involved and the consequential casualties. The statistics relate only to personal injury accidents on public roads that are reported to the police, and subsequently recorded, using the STATS19 accident reporting form. Information on damage-only accidents, with no human casualties or accidents on private roads or car parks are not included in this data. The data include geographical coordinates and exact time (rounded up to the minute level) of the accident occurrence. Using highly detailed GIS maps of the Ordnance Survey (OS VectorMapTM District), I was able to identify the side of each two-way highway that each accident occurred. Using the level of detail of these two datasets, I have matched the information of the two datasets for the whole highway network of England.

Figure 1 displays the mean traffic flow for different times of day for the Leeds area. The traffic

²Over a 15-minute period.

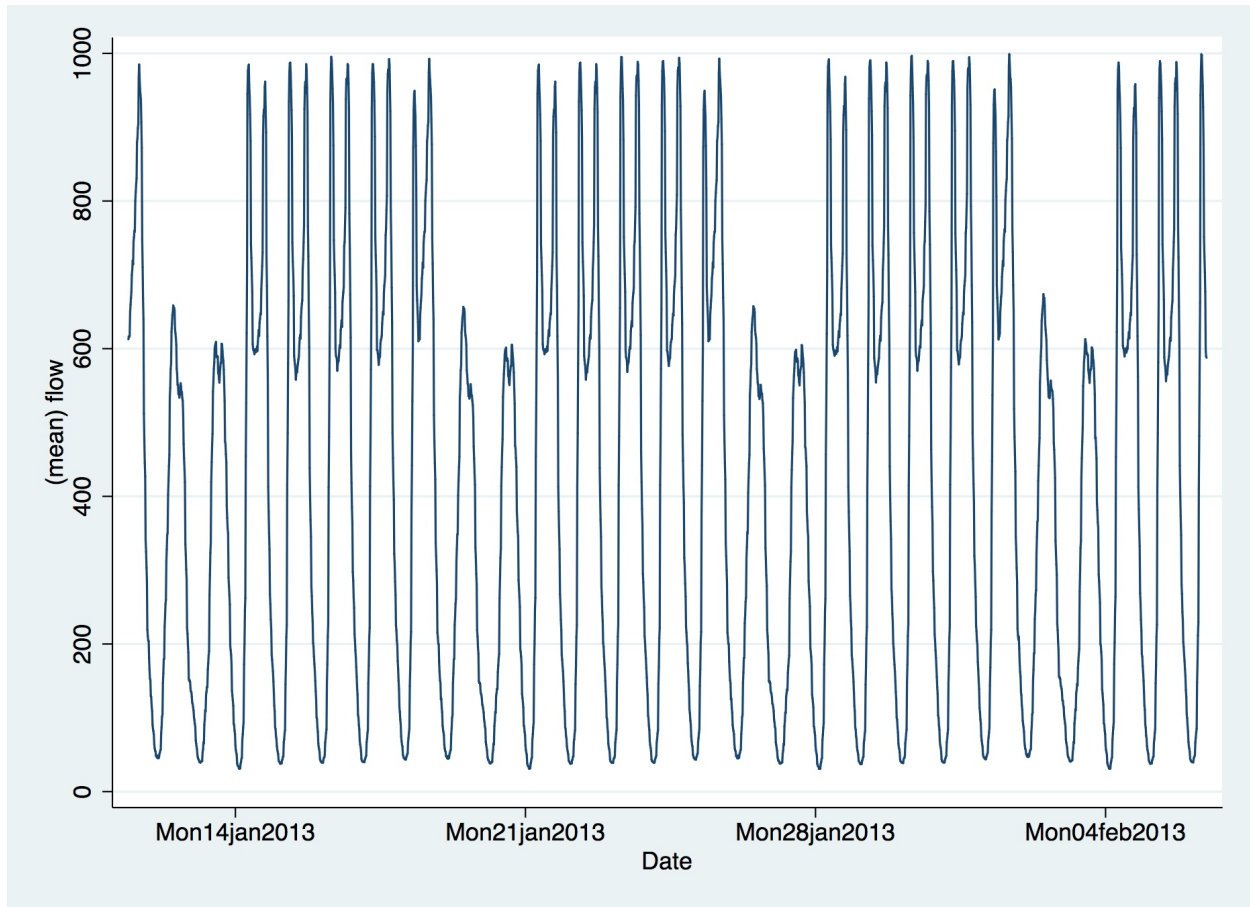
³In this extended abstract version, the data that have been used only include the Leeds area for 2013.

⁴An average of the observed flow for the link, time period and day type.

⁵Please note that journey times are derived from real vehicle observations and imputed using adjacent time periods or the same time period on different days.

flow and average speed data exhibit a remarkably stable periodic pattern, which is repeated every week. These cycles of the traffic flow indicate that out of all the factors that may predict highway traffic, the time of the day and the day of the week are the two most important ones. Using the explanatory power of these two variables, I can observe the “business-as-usual” traffic flow and average speed for a given time period which is roughly unchanged for the whole month or even for the whole year if nothing unexpected happens.

Figure 1: Mean flow for Leeds area highways.



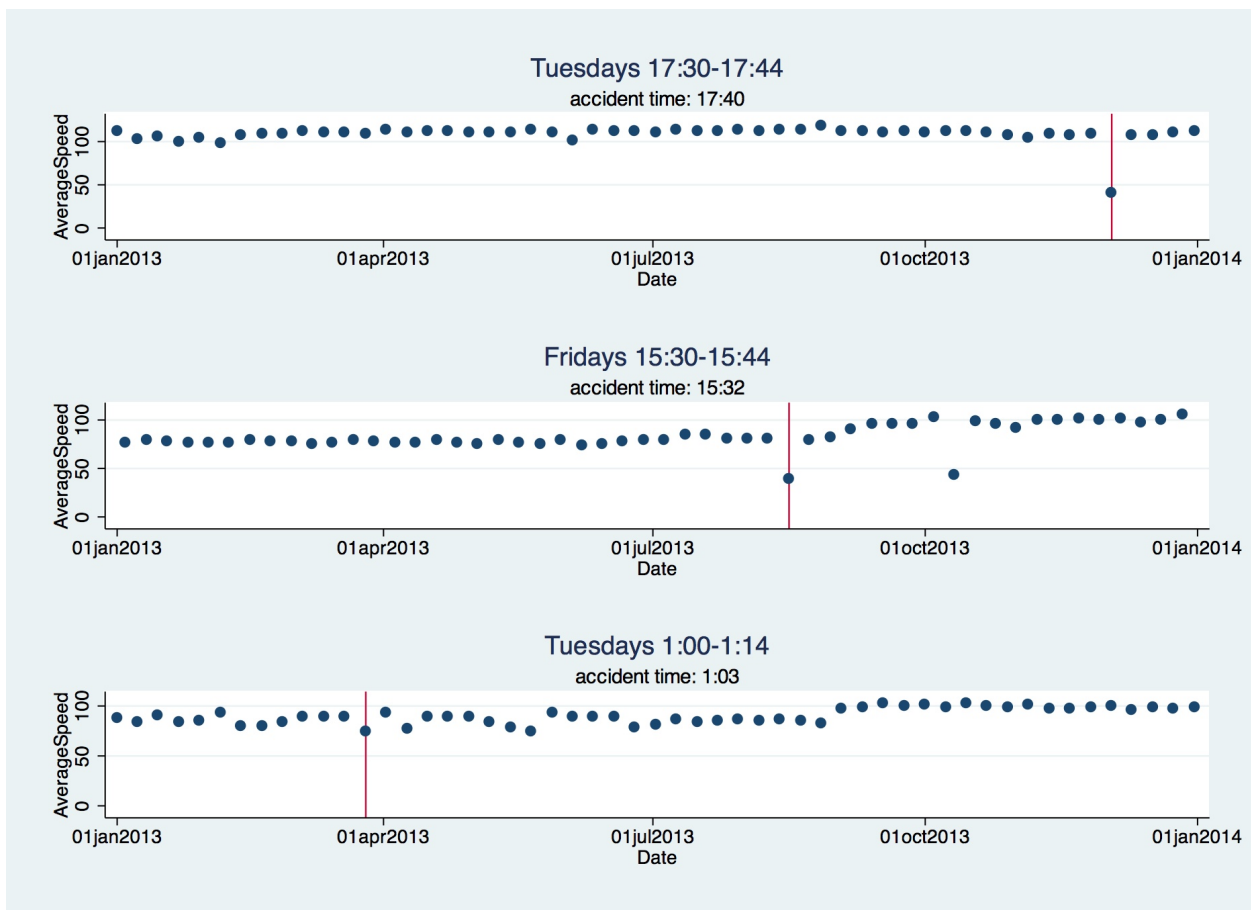
3 Methodology

As in hourly electricity prices in day-ahead markets, traffic flows and average speeds exhibit specific characteristics such as mean-reversion, seasonality and spikes. However, in contrast with electricity markets, traffic flows do not have such a complex time-varying volatility structure. On the contrary, the stable weekly cycles of the traffic flows are those that enable me to estimate the causal effect of highway accidents on average speeds and traffic flows. Nonetheless, this particular characteristic shows that a forecasting model of traffic flows “cannot treat time as one-dimensional”. Time-series models assume that the information set is updated by moving from one observation to the next in time. However, due to the nature of the road travel demand, we adopt the framework proposed by Huissman et al. (2007), which, in this context, treats the 96 time periods of the day (of

15 minutes each) as 96 cross-sectional units that vary from day to day and from highway segment to highway segment.

If an accident happens, we expect that this stable day and time-specific pattern of traffic flow will be disrupted. By being able to observe an almost “perfect counterfactual” of accident absence, the estimation of an accident incidence on traffic flow and average speeds will have a causal interpretation. In figure 2, I chose 3 examples of different times of the day for different highway segments in order to show how the average speed that is observed every week on the same day of the week at the same time is virtually the same for the whole year. In addition, it can be observed that the average speed drops significantly only during the day and the time that the accident happens.

Figure 2: Examples of average speed recurrent stability.

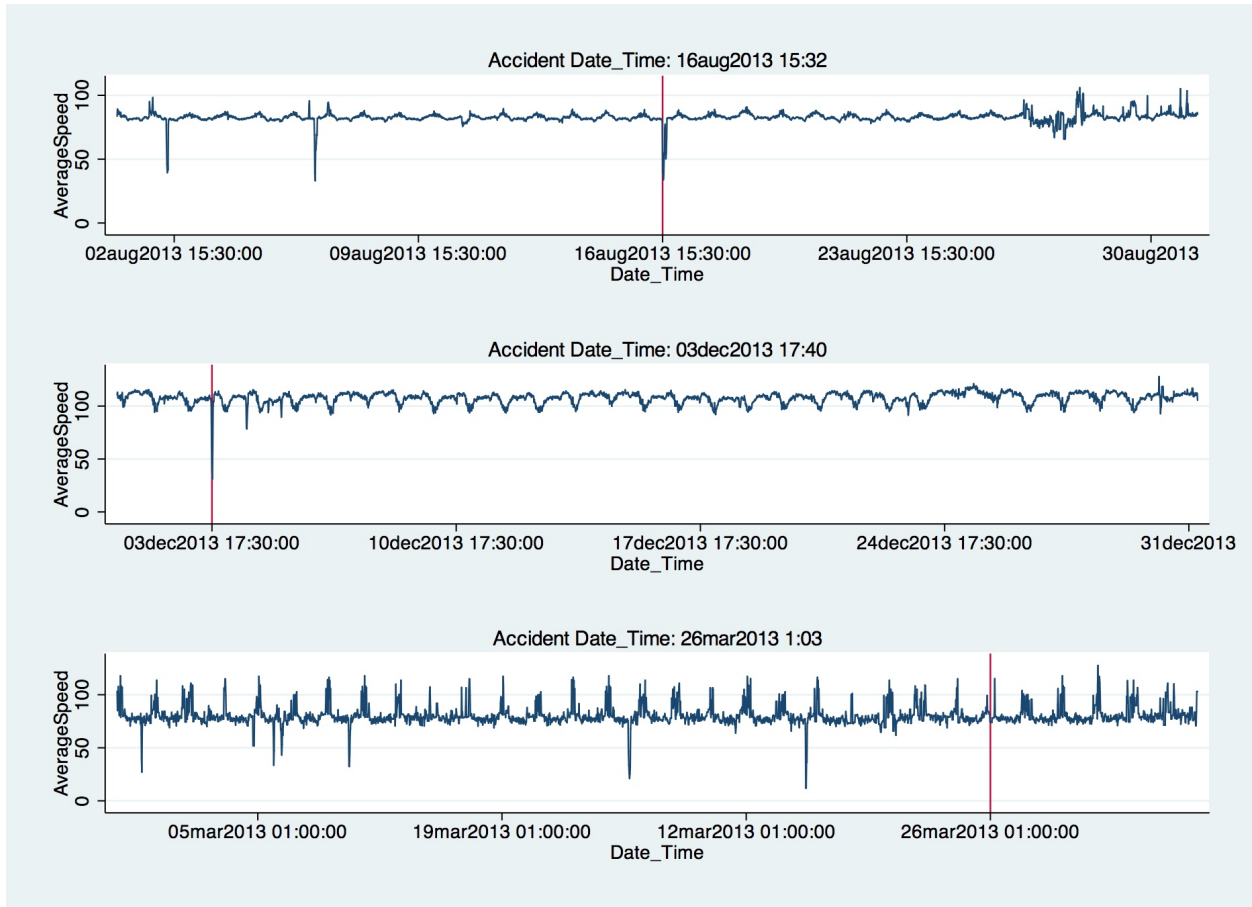


This stability of the average speed holds for most parts of each day. However, during night time and during weekends, this stability is more volatile. This can be explained by the nature of the demand for car travel. Car travel demand is inelastic during the “office hours” of the weekdays (mainly for commuting reasons). This makes the traffic flows, and consequently the average speeds particularly stable during these hours. On the other hand, the last graph of figure 2 shows the average speed during night time (1 a.m.). Even though the average speeds are less stable during this time, we can still observe a lot of stability and a notable decrease of the average speed at

the date of the accident.

Until this point, I have highlighted the persistence of traffic at each particular time of every day of the week. Even if this is true, it should also be mentioned that as most time series processes, traffic flows and speeds at each time of the day also depend on the traffic of the preceding time period (see figure 3). Especially when it comes to highway traffic, one can think of traffic flow dynamics as described by a bottleneck model (for more details, see Small and Verhoef, (2007)). Such a model makes clear why prior traffic matters. This is why ultimately in the simple forecasting model used in this paper, I use both the lagged time period average speed and the average speed during the last week at the same time period.

Figure 3: Examples of average speed's over time variation.



3.1 A simple model

A simple toy model of what will be estimated in this work is presented in the following equation.

$$\ln(speed_{i,d,t}) = \alpha_0 + \alpha_1 \ln(speed_{i,d,t-1}) + \alpha_2 \ln(speed_{i,d-7,t}) + \alpha_3 accident_{i,d,t} + \eta^i + \eta^t * \eta^{day} + \eta^d + \epsilon_{i,d,t}$$

where $\ln(speed_{i,d,t})$ is the logarithm of the average speed in the highway segment i , on the date d and during the 15-minute period t . The logarithms of the lagged average speed vari-

ables for the previous time period ($\ln(speed_{i,d,t-1})$) and for the same time period one week before ($\ln(speed_{i,d-7,t})$) are the main "predictive variables". The dummy variable $accident_{i,d,t}$ takes the value 1 only when an accident has occurred at the highway segment i on the date d and during the 15-minute period t and it is 0 elsewhere. This is the main variable of interest and its coefficient α_3 captures the marginal effect of the accident occurrence on average speeds. Finally, η^i , η^t , η^{day} and η^{date} are highway segment, time period, day of the week and date fixed effects, respectively. Thee time period and day of the week dummies are interacted in order to control for thee day-of-the-week specific time-period trend. Using these fixed effects, I control for observable and unobservable factors that do not change for the same highway segment over time (e.g. road width, quality, security), for variables that do not change over the same time of each weekday (e.g. rush hours, weekends) and for variables that are date specific (e.g. weather, holidays).

An example of an OLS estimation output of such a model is presented in table 1. Table 1 shows that both time period and week lags are highly significant and positive. These two variables are crucial in explaining average speed. The coefficient of the accident dummy implies that the occurrence of an additional accident caused a 12.1% decrease in average highway speeds. However, this is just an example of a result which should be interpreted with caution.

Table 1: Example results.

Dependent variable:	$\ln(speed_{i,d,t})$
$\ln(speed_{i,d,t-1})$	0.718 ^a (0.009)
$\ln(speed_{i,d-7,t})$	0.132 ^a (0.009)
$accident_{i,d,t}$	-0.121 ^a (0.009)
Highway segment FE	Y
Time period of each weekday FE	Y
Date FE	Y
Observations	1,496,440
R-squared	0.78

Notes: Standard errors clustered by time period are in parentheses. ^a, ^b, and ^c indicates statistical significance at 1, 5, and 10 percent level, respectively.

4 Work in progress

Apparently, the model presented and estimated above might suffer from (downward) bias because of the introduction of the lagged variables. I presented this naive estimation because of time and computational power constraints of a standard commercial software package (Stata). This abstract used a sample of the data only covering the Leeds area and for 2013. However, the final estimation will be for the whole of England for the period 2007-2013. In addition, the two effects will be estimated for the different regions of England separately to test for heterogeneous effects. The model which will finally be estimated will include first-differences and it will instrument the first-difference lagged variables with higher order lags using GMM. A similar methodology will be used to estimate the inverse relationship i.e. the effect of traffic congestion on the accident occurrence.

5 Bibliography

Adler, M. W., Ommeren, J. van and Rietveld, P. (2013). Road congestion and incident duration. *Economics of Transportation*, 2(4), 109118.

Arnott, R. and Small, K. (1994). The Economics of Traffic Congestion. *American Scientist*, 82(5), 446-455.

COWI. (2006) Cost-benefit assessment and prioritisation of vehicle safety technologies. Final report. Contract TREN/A1/56-2004. European Commission, Brussels.

Dickerson, A., Peirson, J., and Vickerman, R. (2000). "Road accidents and traffic flows: An econometric investigation." *Economica*, 67(265), 101-121.

Downs, A. (2005). *Still Stuck in Traffic: Coping with Peak-hour Traffic Congestion*. Brookings Institution Press.

European Commission (2011), Impact Assessment accompanying the White Paper, Appendix 5, SEC(2011) 358.

European Commission (2001): White Paper, European transport policy for 2010, time to decide.

Golob, T.F., Recker, W.W., 2003. Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions. *Journal of Transportation Engineering* 129 (4), 342-353.

Huisman, R., Huurman, C. and Mahieu, R. (2007). Hourly electricity prices in day-ahead markets. *Energy Economics*, 29(2), 240-248.

Jansson, J. O. (1994). Accident Externality Charges. *Journal of Transport Economics and Policy*, 28, 31-43.

Jones-Lee, M. W. (1990). The value of transport safety. *Oxford Review of Economic Policy*, 6, 39-60.

Newbery, D. (1987). Road user charges in Britain. Discussion paper 174, Centre for Economic Policy Research (LSE), April.

Newbery, D. (1988). Road user charges in Britain. *Economic Journal*, 98, 161-76.

Noland, R. B. and Quddus, M. A. (2005). "Congestion and safety: A spatial analysis of Lon-

don." *Transp. Res., Part A: Policy Pract.*, 39(7-9), 737-754.

Quddus, M. A., Wang, C., and Ison, S. G. (2010). Road Traffic Congestion and Crash Severity: Econometric Analysis Using Ordered Response Models. *Journal of Transportation Engineering*, 136(5), 424-435.

Shefer, D. and Rietveld, P. (1997). Congestion and Safety on Highways: Towards an Analytical Model. *Urban Studies (Routledge)*, 34(4), 679-692.

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

Snelder, M., Bakri, T. and van Arem, B. (2013). Delays Caused by Incidents. *Transportation Research Record: Journal of the Transportation Research Board*, 2333, 1-8.

Vickrey, W. (1968). Automobile accidents, tort law, externalities and insurance: an economist's critique. *Law and Contemporary Problems*, 33, 464-84.

Vickrey, W. (1969). Congestion theory and transport investment. *American Economic Review*, 59, 251-60.

Vitaliano, D. F. and Held, J. (1991). Road accident external effects: an empirical assessment. *Applied Economics*, 23, 373-8.