The contribution of the paper is an original speed/density model of road congestion costs, and an original model of the economic cost of roadworks and it is shown how these models can be estimated on an extensive and previously unused data set. The present paper reviews the literature and develops these two new models. The initial review is important as it is argued that the existing literature contains a fundamental error in estimating the marginal external congestion cost.

Estimation of these models allow investigation for inner London data of the frequency of hypercongestion, the impact of the proposed new methods of calculating the marginal external congestion costs, the optimality of the charging in the London Congestion Charging Scheme (LCCS) and the impact of roadworks on traffic speeds which has been suggested as the cause of the decline in traffic speeds in the scheme area.

1. **Simple Analysis of the Market Failure of Road Traffic Congestion**

Road congestion can be analyzed through consideration of the road network as a common property resource. The two main characteristics of a common property resource are the use of the resource is restricted to a number of agents and the use by one agent can affect the use by others. We assume that all road users are identical.
and, thus, face the same average journey times. Therefore, an additional road user faces the average journey time and their presence on the road network reduces the speed of all other users and increases their journey times. Consequently, this additional road user imposes an externality on other road users over and above the time costs the additional user faces. (Pigou,1920, Coase,1960, and Polak and Heertje,2000)

In economics, the existence of negative externality causes the market price mechanism to fail in achieving its efficiency. An external cost turns up whenever any consumption or production of goods has a negative impact on others who are not directly involved through the market in that activity. Total social cost takes into account both private and external costs. Because one who causes such external cost pays only his private cost, consequently, he will consume or produce as much as they wish and this leads to either overproduction or overconsumption. Thus the output level is determined by condition that marginal private cost equals to price. Pigou (1920) therefore, proposes to internalize the external cost by charging it as a tax to whoever causes it. This tax should be equal to the optimal marginal external cost. As a result this taxation will adjust market to its optimal level. Concerning road transport, Verhoef (1999) mentions that road use causes a variety of externalities. He distinguishes external costs of the use of road into two types, intra-sectoral externalities and inter-sectoral externalities. Road congestion is suggested as an intra-sectoral externality, with external costs imposed upon one-another by road users, whereas environmental externality, i.e. noise annoyance, is a type of inter-sectoral externality. Therefore the following analysis will discuss only the impact of intra-sectoral externalities, ie road congestion.
The market failure of traffic congestion is demonstrated in Figure 1, we assume identical road users using a uniform road network at a certain time of day. The horizontal axis indicates traffic flow, number of vehicles are passing through a definite length of road in an hour time (vehicles/hour), while the vertical axis specifies costs of using the road. The marginal private cost (\(MPC\)) curve is the unregulated cost of each vehicle per hour (\(£/\text{vehicle/hour}\)) in using a road i.e. fuel costs, driver and passenger time, and vehicles' repair and maintenance costs. Because a road is a common property resource every road users faces a same average social cost (\(ASC\)), which is assumed to be equal to \(MPC\). At low flows, road users can travel at the free-flow speed, and the \(MPC\) is constant. After the traffic congestion develops at higher flows that cause decreases in speed, the \(MPC\) slopes upwards.

Adding the demand curve to Figure 1 represents the marginal private benefit of traffic, the marginal private benefit (\(MPB\)) is assumed to be the marginal social benefit (\(MSB\)) on the demand curve. The marginal social cost curve (\(MSC\)) takes into account of the congestion cost imposed on others by the last road user. After traffic volume is greater than \(F_f\), \(MSC\) diverts from \(MPC\); \(F_f\) is defined as a traffic free flow volume. The difference between the \(MSC\) and \(MPC\) is the marginal external congestion cost (\(MECC=MSC-MPC\)).

The unregulated equilibrium occurs at the intersection of \(MPC\) and \(MSB\), resulting in an equilibrium price at \(P_0\) and at flow of \(F_0\). At this price, an additional road user enjoys his journey but only faces the \(ASC\), whereas other users have to bear an extra cost (\(MECC\)) in term of their extra journey time on the road.
Figure 1: Market Failure of Traffic Congestion

The area under the demand curve represents total social benefits, while the area under the $MPC$ curve represents social costs. Thus, the optimum point is at $(p^*, F^*)$ and the net social cost of unregulated use of the road network is the shaded area, a triangle $ABC$. For this reason, Pigou (1920) suggests internalizing $MECC$ as a tax on every road user, called “Pigouvian tax”. Every road user will be charged with marginal external cost of congestion—the cost which each road user imposes on others on the same road. Hence, under road pricing regulation, every road user is forced to pay the marginal price cost ($MPC$) that is achieved by adding a tax of congestion cost ($MCC$) to the ASC. As a
result, the social optimal is found at the intersection of $MSC$ and $MSB$, and an equilibrium price is at $P^*$ and at flow of $F^*$. The number of vehicles $(F_0 - F^*)$, who suffer from a higher cost of travelling, because their benefit of road usage $(MPB)$ falls below $MSC$, therefore they stop using this road.

Glaister (1981) explains the cost of using the road by supposing $AC(F)$ is an average cost per km to each vehicle, where $F$ is a number of vehicles in a flow of traffic. This cost is assumed to be composed of money costs and time costs. The costs are varied upon traffic speed and on the level of congestion. The total social cost is given by

$$TC(F) = F \cdot AC(F) \quad ; \quad \frac{dAC}{dF} > 0$$

The marginal social cost will be given by

$$MSC = \frac{dTC(F)}{dF} = AC(F) + \frac{dAC(F)}{dF} \quad (2)$$

Therefore, when the impact of road congestion externalities are taken into account, the $MSC$ will be higher than $AC(F)$ by amount of $\frac{dAC(F)}{dF}$; the marginal external congestion cost.

2. Homogeneity of roads and vehicles

The previous discussed studies assume that the road network is uniform. Knight (1924) was the first economist to analyze marginal external congestion costs when the quality of roads differ. He uses Pigou (1920) example of two roads to explain about the congestion tax. The first road is broad without crowding traffic, but is poor graded and surfaced, whereas the second has much better quality, but limited in capacity. The
example assumes many trucks’ drivers, who can freely choose between these two roads. Drivers’ decisions of choosing either of these roads are based on the costs and indeed traffic level is determined by the condition that marginal private cost equals to price. Drivers are more likely to choose the narrow road than the broad road; this would cause more congestion on the narrow road than the broad. However, after the road becomes congested, an additional truck causes external cost on other trucks on the road. Under unregulated condition, an additional truck considers only his marginal private cost. That causes number of trucks on the narrow road greater than the level that is social optimal. These trucks’ drivers will decide to use the broad road when they find no difference of cost and benefit between using either of these roads. Therefore, Knight (1924) suggests to impose a tax on each truck uses the narrow road for trucks’ drivers to bear the marginal external cost of using that road. Walters (1954) also supports Knight (1924)’s idea of non-homogenous roads by proposing that taxation should be varied by different types of vehicles; the slow wide vehicles cause more congestion than the fast narrow vehicles, by different times; volume of vehicles are different by different time of the day and by different roads. Although, Walters (1961) argues that an assumption of traffic homogeneous is unrealistic, he agrees that its simplicity is useful for theoretical development and empirical testing. Accordingly, his study presumes sameness in all vehicles, drivers, owners, and also each vehicle will definitely face the same speed and cost at a given volume of traffic. This assumption of Walters (1961)’s is developed from Pigou (1920), Knight (1924),and Beckman, McGuire and Winster(1955) and Walters (1954).
In addition, Walters (1961) also develops Pigou (1920)’s two roads case which are connecting a same starting and destination point to explain an efficient distribution of traffic between these roads by using a supply-demand for traffic flow diagram in Figure 2. This diagram defines $AC_1$ as the private unit cost of a vehicle using the first road and $AC_2$ is the private unit cost of a vehicle using the second road. Additionally, roads are perfect substitutes for one another; therefore there is only one demand curve. Both cost curves initially are constant at their free flow speeds and, after congestion starts, the cost curves are positively upward sloped which implies rising in unit costs because of congestion.

At lower flows, the unit cost on the second road is lower than the unit cost on the first road. A large number of vehicles are using these roads and each of them is making decisions only taking into account his private cost. No one will worry with the congestion cost that he imposes to other vehicles. $MC_1$ and $MC_2$ are marginal private costs are measured from private cost curves $AC_1$ and $AC_2$ respectively. $AC_T$ is traffic flow by horizontally summing of private unit cost curves. Walters (1961) makes a comment about the condition of efficient distribution and Beckman, McGuire, and Winsten(1955) indicate that it will occur where marginal cost $MC_1$ at $F_1$ is equal to marginal cost $MC_2$ at $F_2$. But Walters (1961) argues that it should occur when total traffic flow reduces from $F^*$ to $F_E$, indeed at the intersection where aggregated marginal social cost $MC_T$ equals to $PF_E$. 
Interestingly, Walters (1961) examines road congestion at a bottleneck section. He explains that traffic flow with the higher demand is lower than traffic flow with the lower demand at the bottleneck.

Although a pure bottleneck type does not represent a general congested situation, Vickrey (1969) still suggests that it is useful for a simple analysis because it provides some valuable understanding into the nature of overall traffic congestion problem. He explains vehicles’ decisions in travelling through a bottleneck. If during a given period; demand for travelling exceeds the bottleneck capacity, queuing occurs. As a result,
some vehicles will have to arrive at the destination either early or late. Therefore, vehicles’ decisions depend to degrees of queue, time spending in queue has to be taken into their decision, vehicles’ time value is an additional factor for decision making. Vickrey (1969)’s example gives a uniform value of time for all commuters spent, at home, at the destination, in the queue, and at the destination after desired time.

3. The Road Traffic Speed-Flow Relationship

In this section, we consider studies which use the road traffic speed-flow relationship to examine the economics of road congestion. This traditional approach is the most frequently used method to examine the economics of road congestion. Walters (1961) shows how it is related to the road traffic speed-density approach and the advantage of this latter approach. An increase in number of vehicle entering a road increases density, and this reduces an average speed. The conventional concept describes that an increase in the traffic flow reduces average speed where the effect on vehicle kilometres of an additional vehicle is greater than effect of average speed reduction. However, a decrease in traffic flow could reduce average speed after an additional vehicle enters the road after traffic flow has reached the road capacity. (Walters 1961, Button, 1993, Verhoef, 2003 and et al.)

Figure 3 demonstrates the speed-flow relationship on an uniform road over a period of time at different speed levels. The vertical axis indicates speed and the horizontal axis indicates flow. The diagram explains that at low traffic flows $F_A$ allows vehicles travel at a high speed at $A$. But as more vehicles enter the road, they will slow other vehicles’ speed down and, therefore, at a higher traffic flow $F_B$ the average speed falls to $B$. As
long as the effect of additional vehicles override the reduction in average speed, the flow is increasing until the effect of additional vehicles offset the decrease of average speed. At that point, the maximum flow $F_M$ takes place at the road capacity. Newbery (1988) adds an important notion that $F_M$ represents a boundary between free (at point $D$) and congested flow (at point $C$) and he makes a further reference of Hall et al. (1986) that whenever inflows fall below $F_M$ the traffic will return to free-flow conditions. However, vehicles may still try to enter the road beyond the maximum flow level, because they do not have enough information. This occurrence causes more speed reduction at point $D$ and $E$ while traffic flow is reducing from $F_B$ to $F_A$ respectively, indeed the speed-flow curve bends backward.

![Figure 3 Speed-Flow relationship](image)

**Figure 3 Speed-Flow relationship**

Newbery (1988) also analyzes the speed-flow relationship in Figure 3 but he assumes that vehicles are travelling through a bottleneck, and increasing traffic flows through
bottleneck reduces speed from A to B. However, once traffic flow reaches its maximum capacity at the bottleneck capacity, the condition of the speed-flow relationship is changed. Then falling flow will decrease an average speed. Quinet and Vickerman (2004) define that the upper part of curve represents a stable flow, and the lower part , backward-bending curve represent an unstable flow. Interestingly, Newbery (1988) explains an unstable lower curve for a bottleneck part of road. He describes that a decrease in traffic flow makes the rate of traffic leaving the bottleneck faster than the arrival rate, and the average speed possibly jumps back to A at the same flow. However, he comments that this relation is only for road design, not for estimating the cost of congestion. He mentions that the cost of congestion estimation is based upon the total extra time which the remaining traffic takes to complete their journeys and not their speed at a particular point on the road network.

Hence, Newbery (1988) suggested the marginal time cost (MTC) as a better measuring of speed-flow relationships. He explains that it is only estimated from the downward-sloping part, the upper curve. But this measurement eliminates the backward-bending section. The analysis of MTC has been found in Harrison et al. (1986), the modeling of traffic flow in Hongkong. In addition, Newbery (1988) mentions the analyses which were presented by Duncan (1980) , Harrison (1986) and by Department of Transport in COBA-9 (DTp,1987) or QUADRO-2 (DTp,1982). He suggests that the relationship of speed-flow is linear. He also suggests that the linear relationship is convenient for measuring the marginal external congestion cost (MECC) and also the marginal time cost (MTC).

Giving the travel cost per kilometre of a vehicle is

\[ AC = a + \frac{b}{V} \]  \hspace{1cm} (3)

where \( \frac{b}{V} \) is the cost of congestion per vehicle hour, including the opportunity cost of the time of the vehicle therefore the total cost of a flow of \( q \) vehicles per hour is \( TC = AC \cdot F \). If an additional vehicle is added to the flow, the total social cost is increased by;

\[ \frac{dT C}{dF} = \frac{dAC \cdot F}{dF} = AC + F \frac{dAC}{dF} \]  \hspace{1cm} (4)

From (4), the first term is the private cost borne by each vehicle, and the second is MECC borne by other vehicles. From (4), MECC is given by

\[ MECC = \frac{b}{V} - \frac{dlnV}{dlnF} \]  \hspace{1cm} (5)

4. The Road Traffic Flow-Density Relationship

According to Walters’ (1961) observation, he makes note that an increase traffic demand does not always increase traffic flow. At a very high traffic demand, the traffic flow is lower than at a lower demand. In spite of traffic flow, he proposes traffic density because traffic demand directly depends on traffic density. Haight (1963) explains that traffic flow depends on traffic density because the flow is the product of density and
speed. Figure 4 presents the traffic flow-density relationships. The vertical axis indicates traffic flow, and the horizontal axis indicates traffic density. The curve is starting with no vehicle on the road, zero density, this also gives a zero traffic flow. While vehicles drive at a free flow speed level; an increase in number of vehicles does not much affect speed reduction, more vehicles entering to the road results gradually increase in traffic density that also dramatically increase traffic flow. Once traffic becomes congested, an increase density will steadily slow down an increasing rate of flow until the flow is maximized at the road capacity. After that an increase density will reduce flow. Until the density reaches the maximum level $D_M$, no more vehicle can enter to the road, traffic flow is stopped at a zero speed.

Figure 4 Traffic Flow-Density Relationship
5. Hypercongestion

According to Newbery (1989), the conventional estimation of MECC ignores the backward-bending positive sloped part of the speed-flow curve for the purpose of convenience in measurement. The backward bending curve is known in economics as hypercongestion (e.g. Lindsey and Verhoef, 1999 and Small and Chu, 2003). In Newbery’s (1988) analysis, he suggests that the backward bending curve represents unstable equilibrium as there is no unique speed corresponding to each flow level. Verhoef (2003) clarifies that economic models generally analyze traffic congestion cost under assumptions of constant speed, flow, and density along a given single road over time. However, constancy of these variables over time only holds for stationary equilibrium. Conversely, a dynamic analysis is required for when these variables vary over time.

Lindsey and Verhoef (2003) use Figure 5 to examine the road traffic speed-density-flow relationship. The diagram is developed from an identity in which traffic flow \( F \) is a product of speed \( S \) and density \( D \). They assume traffic density is an average quantity of vehicles on a given road, whereas speed is an average journey distance in an hour. As a result, a unit of traffic flow is an average rate of vehicles-kilometres flowing in an hour.
Figure 5 The Road Traffic Speed-Density-Flow Relationship Diagram

Figure 5 explains speed-flow-density relationships. The first quadrant shows how an increase in density decreases average speed, the second quadrant shows how traffic flow is related to speed (and is the same as Figure 3). The fourth quadrant represents how flow and density are related.

The following discussion will consider the second quadrant as a result of the first quadrant. Starting at the first quadrant, the horizontal axis denotes traffic density and the vertical axis denotes traffic speed. Whereas, in the second quadrant, the vertical axis still denotes traffic speed and the horizontal denotes traffic flow. The road traffic speed-density-flow relationship is described into four possible scenarios, distinguishing by level of traffic density.
$D, V,$ and $F$ symbolize densities, speeds, and quantities of vehicles sequentially.

1st: if $0 \leq D \leq D_L$, then $V_f \geq V \geq V_L$ and $0 \leq F \leq F_1$;

Firstly, where traffic density $D$ is increasing in the low range level from zero to $D_L$ vehicles, all road users enjoy driving at high speed: free flow speeds vary from $V_f$ to $V_L$. As a result, the rate of traffic flow greatly increases from zero to $F_1$.

2nd: if $D_L \leq D \leq D_0$, then $V_L \geq V \geq V_0$ and $F_1 \leq F \leq F_M$;

Secondly, the traffic speed slightly decreases from $V_L$ to $V_0$ against an increase in density from $D_L$ to $D_0$, meanwhile, the traffic flow slowly increases from $F_1$ to $F_M$.

In the first two scenarios, increasing density has more of an impact than decreasing speed and, hence, flow still increases as density increases.

3rd: if $D_0 < D < D_H$, then $V_0 \geq V \geq V_H$ and $F_M \geq F \geq F_1$;

Thirdly, the traffic speed dramatically decreases from $V_0$ to $V_H$ against an increase in density from $D_0$ to $D_H$, simultaneously traffic flow decreases from $F_M$ to $F_1$. This implies an increase in density has a smaller impact on flow than the decrease in speed and, therefore, flow starts to decline.

4th: if $D_H < D < D_M$, then $V_H \geq V \geq 0$ and $F_1 \geq F \geq 0$;

Finally, the fourth scenario, as traffic density approach road capacity, the impact of an increase in density over the range from $D_H$ to $D_M$ has smaller impact on flow than the striking decrease in traffic speed, which fall from $V_H$ to zero, synchronously, shrinking
traffic flow from $F_1$ to zero. At the maximum traffic density $D_M$, both traffic speed and traffic flow become zero, which is to say the traffic has completely stopped.

From the above diagram analysis, Lindsey and Verhoef (2003) address two meaningful economic implications. The first phenomenon is called the congestion when traffic flows increase from zero to its maximum ($0 \leq F \leq F_M$). The second phenomenon is called hypercongestion when traffic flows decrease from its maximum to zero ($F_M \geq F \geq 0$).

Therefore, the maximum level of traffic flow can be seen as a border line between a congested and hypercongested phases, which has significant implications in economic study. Likewise, Johnson (1964) applies the speed-flow-density relationships to estimate time congestion costs. He also concludes with two alternatives types of traffic congestion. The first demonstrates rising average time-cost with rising flow, or the speed-flow relationship with a positive slope, this is called ordinary congestion; and the second demonstrates rising average time-cost with falling flow, or the speed-flow relationship with a negative slope, this is called hypercongestion. In addition, Walters (1961) suggests that on the backward sloping part of time-flow cost curve, the marginal social cost cannot be measured on time-flow cost curve because the change in flow is negative. This is also suggested by, for example, Newbery (1990), and Small and Chu (2003) that hypercongestion equilibrium is inefficient.

Evan (1992) discusses Johnson’s idea (1964) that proposes to analyze congestion cost by using density instead of flow but does not formally analyze congestion using a speed density approach.
Ohta (2001) also recommends using density rather than the flow in considering congestion costs but does not formally consider the idea of marginal external congestion cost. Verhoef (2003) also develops a congestion cost model by using a speed-density relationship but then transforms this into a speed-flow relationship.

Many economists (e.g. Chu and Small, 1997, Lindsey and Verhoef, 1999,2000, Verhoef, 1997,2003, and Kuwahara,2007) discuss but do not solve the problem of the inadequacy of speed-flow relationships in estimating hypercongestion costs. In addition, it is notable that low speeds and low flows of hypercongestion are very common in urban area. (Lindsey and Verhoef,1999 and Small and Chu, 2003). They claim that the inefficiency is caused by an unrealistic assumption of uniformity of speeds, flows and densities along the road. For example, Lindsey and Verhoef (1999) comment that hypercongestion usually occurs on nonuniform roads, i.e. in upstream queues of a saturated bottleneck. Likewise, Small and Chu (2003) argue that where quantity demand and supply are represented by traffic inflow and outflow respectively, and assume that the inflow or outflow are always constant along the road. As a result, an assumed condition of uniformity of speeds, flows and densities along the road is invalid. This argument is also agreed with by Kuwahara (2006) who explains that the development of a traffic queue occurs when the demand exceeds road capacity. Therefore, on the road network there are differences in speeds, flows and densities. Therefore, a dynamic solution is required to model this phenomenon. Lindsey and Verhoef (1999) propose that the analysis of congestion costs distinguishes static modeling and dynamic modeling.
In order to distinguish analyses of static and dynamic cost congestion modeling, Verhoef (1999) suggests to differentiate two basic types of demand; peak demand for dynamic analysis and continuous demand for static analysis. He defines a peak demand as a case where a limited number of potential users consider using the road during the same peak period. Whereas, a continuous demand is referred to as a case of constant demand over time. Verhoef (1998, and 2003) defines a static model as a time independent model that assumes constant variables over time with respect to traffic speeds, flows, and densities. Whereas, a dynamic model is defined as time dependent and the aforementioned variables vary over time.

Lindsey and Verhoef (1999) indicate that studies of road congestion costs are classified into two levels. The first is the individual vehicles interaction level, and the second is the aggregation level; treating vehicle as a fluid-like band.

Empirically, most researches are done in the individual vehicles interaction level. For example, Verhoef (2003) develops congestion cost modeling using car-following theory in a network to analyze the speed-density relationship and transform into the speed-flow relationship. (this study is based on an original model developed by Verhoef, 2001). However, though this model considers a speed-density relation, it does not consider the marginal external congestion cost.
6. Why should the speed-density modelling be used for the estimation of marginal external congestion cost?

A speed-flow relationship is the conventional approach to the estimation of the marginal external congestion cost. This relationship can be divided into two non-linear parts. The first part shows ordinary congestion with a negative speed-flow relationship, whereas the second part shows hypercongestion with a positive relationship. There is a turning point, where the speed-flow relationship switches from a negative to a positive relation. This point indicates the maximum flow level or the engineering road capacity. The empirical estimation of a speed flow relation is difficult as the relation from flow to speed is not function and the relation is non-linear. In particular, the estimation of the turning point from ordinary to hypercongestion is likely to be difficult. Most empirical studies consider only negative speed-flow relationships, whereas in urban areas traffic is likely to be hypercongested frequently. In many urban areas, grid lock or very low traffic speeds throughout the day are common, e.g. London, Bangkok and Manila. It is shown in the later following analysis that marginal external congestion cost can only be properly and accurately explained in the context of a speed-density relationship and, in the case of hypercongestion, it cannot be explained by the traditional speed-flow relation. Thus, for important theoretical and empirical reasons congestion must be investigated using a speed-density relationship.

We use Figure 6 to explain the speed-density-flow relationship. In the first quadrant, the diagram shows a non-linear speed-flow relationship with the y-axis representing traffic speed and the x-axis representing traffic flow. The second quadrant, presents a speed-
density relationship, where the y-axis represents traffic speeds and the x-axis represents traffic density. The non-linear traffic speed-flow relationship in the first quadrant indicates that each flow level comes with two speeds, i.e. the relation is not a function. The curve is divided into two parts: an upper curve with a negative slope representing ordinary congestion, whereas the lower curve has a positive slope representing hypercongestion.

Figure 6: Speed-Density-Flow Relationship

In the first quadrant, the turning point $F_{max}$ is the road capacity, and at this point, the curve bends backward switching from ordinary congestion to hypercongestion. Providentially, the second quadrant demonstrates a monotonic traffic speed-density curve with a negative slope throughout. In transport economics, a fundamental equation is that traffic flow is a product of traffic density and speed. On the upper part of the first quadrant curve, an increase in traffic density decreases the average traffic speed and increases traffic flow simultaneously. However, on the hypercongestion part of the first quadrant curve.
quadrant curve, an increase in traffic density decreases traffic speed and traffic flow simultaneously. This occurs as the increase in the number of vehicles is accompanied by such a large fall in traffic speed the total number of vehicle-kilometre decreases. This comparison emphasizes that the examining of traffic congestion impact through using traffic density is sounder than traffic flow because an increase in density always decreases speed. This supports the use of the speed-density relationship in a road traffic congestion model and this approach is shown to be valid more formally below.

Suppose $F$ vehicles per hour are flowing through a uniform road at average speed $V$ kilometres per hour. In addition, assume the vehicle average cost per kilometre $AC$ using this road is composed of two costs; $a$ is a fixed private cost, and $b/V$ is a time cost.

\[ AC = a + \frac{b}{V} \]  

(6)

Hence, total cost of using this road in an hour is

\[ TC = F \cdot AC \]  

(7)

Differentiating (7)

\[ \frac{dT C}{dF} = AC - F \cdot \frac{b}{V^2} dV \]  

(8)

Equation (8) shows the marginal social cost of using the road is composed of two components. The first term is the marginal private cost which equals the average cost
and the second term is the marginal external congestion cost \((MECC)\) imposed by an extra vehicle kilometre on other road users. Without congestion, the second term is zero.

Remarkably, equation (8) only holds for ordinary congestion, because of \(dV/dF<0\). Whereas, when hypercongestion occurs \(dV/dF>0\), this implies that the marginal social cost is lower than the average cost. This is odd but what is happening in hypercongestion is that flow increases when the number of vehicles on the road falls. It is not that the extra vehicle kilometres are generating a positive externality (a very odd concept) but that the reduced number of vehicles is giving a positive externality to the remaining road users.

Thus, we continue analyzing the congestion costs using a speed-density relationship. Traffic density \(D\) is the average number of vehicles along a finite length of a uniform road, and traffic flow \(F\), is a number of vehicles in an hour which passes through that road, at a traffic speed \(V\), is an average distance driving in an hour.

An increasing traffic flow \(F\) decreases an average speed \(V\) as follows

\[
V = f(F) : \frac{dV}{dF} < 0 \quad (9)
\]

Computing traffic flow \(F\) as a product between traffic density \(D\) and traffic flow \(F\)

\[
F = DV \quad (10)
\]

A change in traffic flow is composed of two parts as follow

\[
dF = dD.V + dV.D \quad (11)
\]
Equation (11) shows that a change in traffic flow is composed of two parts. The first part represents an increase in flow because of additional vehicles driven at the average speed and the second part presents a decrease in flow because of existing vehicles driving less miles in the given time period as the speed has fallen. Therefore change in traffic flow can be either an increase or a decrease, this depends to whether an impact of additional vehicles is greater than speed reduction or not. The traditional analysis would appear to be wrong in that an one unit increase in traffic flow is composed of one extra vehicle travelling more than one extra kilometre but (in the given time period) all other vehicles travelling less as the traffic speed has decreased. Thus, the marginal external cost in equation (8) has to be considered for more than one extra unit of travel by the additional vehicle. Thus, the traditional analysis has to be reformed to allow for this larger denominator when estimating the marginal external congestion cost.

Similarly, we can use this new approach to consider the marginal external congestion cost in the case of hypercongestion. The literature mostly ignores this important problem. The following analysis using the speed density approach can be applied to both ordinary and hyper congestion. From (10), we replace $F$ in (7)

$$TC = D.V.AC$$  \hspace{1cm} (12)

Then differentiating (12)

$$\frac{dT_C}{dD} = V.AC + D.\frac{dV}{dD}.AC - D.V.\frac{b}{V^2} \frac{dV}{dD}$$  \hspace{1cm} (13)

Equation (13) shows an extra vehicle produces an externality. The marginal social cost for an extra vehicle is composed of three parts. The first term is the marginal private cost which each vehicle has to pay for travelling $V$ kilometres. The second term is a
negative value, because of \( \frac{dV}{dD} < 0 \), it implies a reduced cost where existing vehicles drive less miles as a result of the additional density lowering speed. The last term gives a positive value, it is the marginal external congestion cost. Moreover, equation (13) also shows that the speed-density relationship should be considered in estimating the marginal congestion cost for ordinary congestion as well as for hypercongestion phenomenon.

In view of aforementioned literature, this paper proposes a new contribution in which the marginal external congestion cost of ordinary and hypercongestion should be estimated using speed-density relationships.

**7. The Impact of Roadworks on Congestion Costs**

Roadworks reduce the capacity of a road network and thus affect road works. We are not aware of any substantive economic analysis of the impact of roadworks on congestion costs. This omission is important as road builder contractors have to pay for the suggested costs of roadwork overruns (Transport for London, 2005). Additionally, roadworks are often cited as causes of the reduction of traffic speeds in urban areas (Transport for London, 2005).

Roadworks reduce effective road space and, as a consequence, traffic density is increased. Thus, the traffic congestion causing by roadworks and the general speed-density congestion are interwoven. For instance, roadworks during a low traffic density do not cause as much traffic congestion as during high traffic density. Conversely, a
small level of roadworks can cause a traffic standstill whilst traffic density level is high. Therefore the impact of different levels of roadworks are assumed to shift the traffic speed-density curve downward. Figure 8 demonstrates the speed-density relationships with roadworks.

![Graph showing speed-density relationships with and without roadworks]

**Figure 7 Speed-Density relationships with Roadworks**

The impact of roadworks is to reduce the flow and density of vehicles. Some road users will choose to not drive or take alternative routes and, thus, avoid the increased costs of passing through the roadworks. Thus, an overestimate of the economic costs of the roadworks can be derived from using the pre-roadworks traffic flow and the reduction is traffic speeds from the roadworks. Similarly, an underestimate of the costs can be derived from using the post-roadworks traffic flow which fails to take account of the traffic lost because some road users taking alternative routes or do not make the journey at all.
8. Data Availability

The following paper will discuss the econometric modeling of speed-density and roadworks on a uniform road in the London Congestion Charging Scheme (LCCS). The intention is to use spline estimation or possible alternative methods on speed and density data collected from the LCCS area. This data is composed of three parts:

1. Hourly traffic flows data at 46 counters measured from January to December 2008, supplied by Transport for London (TFL),
2. Speed data measured from January to December 2008 supplied by a private company TrafficLink for 100,000 UK vehicles with GPS monitoring.
3. All Roadwork data inside the M25 area and collected from January to December 2008. The data has been supplied by Traffic Link.

The collection of this data has been problematic but all of (1) has been received, a sample of (2) and most of (3). The intention is to use a representative sample of counter data and match this with the speed and road work data. Protocols for selecting the counters and matching with speed and road work have been established and we await the delivery of the final data.

9. Conclusion

This paper develops new speed/density models of marginal external congestion cost (MECC) for ordinary and hypercongestion and an original model of the economic cost of roadworks. We have clearly shown a fundamental error in the existing literatures which estimate the speed/flow model of marginal external congestion cost. Hence, the paper
presents how these new models can be estimated on new and large urban data set, the estimation of these models is allowed for London Congestion Charging Scheme (LCCS) area. In fact, these models investigate the frequency of hypercongestion for LCCS area, and emphasize a comparison of marginal external congestion costs with LCCS charge. Another cause of concern is the fact that the impact of roadworks on traffic speeds which causes the decline in traffic speeds in LCCS area, therefore the paper includes the investigation of roadworks impact in this area.

References


