CHAPTER 6

CONCLUSIONS

6.1 The Development of a New Model

The literature on the economics of road congestion has seen a number of inconclusive debates on the correct basis for the appropriate analysis. In spite of this unresolved confusion, the traditional speed-flow relationship has been used extensively in the estimation of marginal external congestion costs. The ambitious purpose of this thesis has been to develop an economic model that resolves some of these debates and provides a sound theoretical basis on which to estimate marginal external congestion time costs. In particular, it is important to develop a model that explains the marginal external congestion time costs of hypercongestion and the marginal external congestion time costs of different types of vehicles.

In the new model, the road user is assumed to make a journey along a uniform road that is part of a road network. The user compares journeys with different start (and finish) times. The different journeys involve occupation of the road at different times. At these different times, there will be different road densities and consequently different speeds. The user chooses that journey timing that gives them the greatest consumer surplus. However, this choice will increase density in the chosen consecutive time periods and reduces the speed of other users. In turn, these other users will take longer to complete their journeys (and some, in the longer term, will decide not to make their journeys) and cause more congestion. This pattern of cycles of additional congestion is repeated until there is sufficient additional road capacity to allow journeys to be completed and, in the longer term, some journeys may be deterred. These longer journeys are all external costs and are the result of the increased density. This simple model is used to explain various strongly debated issues in the literature on the economics of road congestion, e.g. the disputes between Else (1981 and 1982) and Nash (1982); Demza and Gould (1987) and Evans (1992a); Evans (1992a and 1992b) and Hills (1993); and Ohta (2001a and 2001b) and Verhoef (2001a). In particular, the focus of the model on road density or occupation allows consideration of the external costs of hypercongestion. Some past studies seem to suggest that the upward sloping speed-flow curve in hypercongestion implies "implausibly negative congestion externalities" (Newbery, 1989). As the speed-density curve is always downward sloping, the model's analysis of hypercongestion always gives an external cost and it can be shown that this is much greater for hypercongestion than ordinary congestion.

The simple model of road user behaviour is used to develop equations that allow a correct estimation of the marginal external congestion time cost. The starting point for this development is that, in one time period, it is an extra vehicle that causes congestion not additional flow. Thus, an extra vehicle leads to additional travel by this vehicle in this time period but all other vehicles travel a little less. Thus, the aggregate change in flow is composed of two parts: an increase given by the extra vehicle and a reduction in distance travelled by all other vehicles. It is important to compare this theoretical point with the actual possible error potentially caused by using the conventional analysis. If densities and speeds are constant, it is the case that the two analyses give identical results for the marginal external congestion time cost. In this case, the theoretical development considered in the thesis is not of significant importance. However, if the densities, flows and speeds vary over time, there can be very important differences and the new model is the appropriate method to examine the variation in marginal external congestion costs across time.

The model also shows that there are two external costs resulting from increased density. The first cost is that of the above discussed greater length of journey time. Secondly, the longer journey results in a demand side externality in that the user has to travel earlier or arrive later and this reduces their consumer surplus.

The analysis shows the relation between the speed elasticities with respect to density and flow and, more importantly the underlying causes of the external effects and costs. These relationships are encapsulated in

$$b \quad \frac{dV}{dD} \quad \frac{D}{V^2} \quad = \quad \frac{b}{V} \quad \frac{dV}{dD} \quad \frac{D}{V}$$

$$= \quad \frac{b}{V} \quad \frac{\varepsilon_{VF}}{1 - \varepsilon_{VF}} \qquad (6.1.1)$$

From such relationship it is possible to estimate the marginal external congestion time cost on observable data.

Additionally, the expression (6.1.1) answers the question of how can hypercongestion can be associated with a negative externality when it is defined by a positive slope in the relation between speed and flow? The simple answer is that it is the slope of speed with respect to density that determines the congestion externality and this is always negative.

It is shown that hypercongestion can represent an equilibrium outcome but this is not efficient. All hypercongestion equilibria are shown to be unstable as any disturbance in density in a given time period has an explosive impact on hypercongestion in the adjacent time periods. This instability results from, in the case of hypercongestion, the elasticity of speed with respect to density exceeding one in magnitude.

The major purpose of this thesis is to develop a theoretically sound model of the economics of road congestion that can be used to estimate marginal external congestion time costs. Thus, the estimation of the three theoretical models developed here are discussed in turn.

6.2 Empirical Analyses

The empirical analysis in this thesis examines the marginal external congestion time cost for ordinary congestion on eight streets in the Central of London for data from late 2008. A variety of different specifications (log, quadratic in logs, linear, quadratic, SPLINE and SPLINE in logs) and one estimation technique (ordinary least squares) are examined. All estimations were subject to the various standard diagnostic statistical tests. The estimates nearly all give negative elasticities of speed with respect to density on each street. The estimates are in the order of 0.1 with a large degree of variation between methods and streets. This suggests that the marginal external congestion time cost is in the region of 10% of the average private time cost. Thus, it is concluded that ordinary congestion only results in a slight to moderate externality. A simulation exercise showed that for periods when density and speed are constant the new and conventional methods, not surprisingly, give the same estimates of the marginal external congestion time cost. However, as soon as variation in density and speeds takes place there are important differences in these external costs.the day

There are four important and closely related issues to be considered in the empirical estimation of the marginal external cost of congestion. Firstly, the selection of data needs to separate out carefully between what is considered to be ordinary congested and hypercongested flow. It is important that a turning point is identified that gives the maximum flow. This can only be achieved through observation of the data and making a judgement that is tested for robustness by examined alternative choices. This has not always something that has been carefully considered as there are many examples of estimation of simple linear specifications.

Secondly, the model of road congestion implies an important non-linear relation between speed and flow and by definition there is a turning point. Thus, it is important to test for the correct non-linear model and it is suggested non-parametric models may be of great use. Fitting LOWESS curves to the underlying data may help in identifying the turning point of the speed-flow relationship.

Thirdly, the issue of the correct approach to estimating the relation between speed and flow or density needs to be considered. The simple identity between flow and the product of speed and density affects how the estimation of the marginal external congestion cost should be approached. It might be inferred from the use of models with speed on the vertical axis or as a left hand side variable that flow is being taken as causing speed. It is suggested that this unexplored subject requires more investigation and has been neglected as a problem in the study of traffic speed and flow.

Flow data is typically measured over a relatively long period and is used in empirical studies as an average or interpolated measure. Consequently, it is likely to have a degree of error in variable measurement. This causes a bias downwards in two variable ordinary least squares. This issue is mostly ignored in the estimation of marginal external congestion time cost.

The issues discussed above apply equally to the case of hypercongestion. It would appear that the present study is the first one to make an empirical investigation of hypercongestion using the speed flow relationship. The simulation exercises show how data and the estimated results for elasticities of speed with respect to density can be used to examine marginal external congestion time costs. These are often substantial in the region of 100 minutes per vkm. The estimates are larger at the beginning of a period of hypercongestion as the increased density expands explosively because of the lack of additional road capacity. Additionally, whilst deterred demand has very little impact on the size of the METC in periods of low congestion, it has very important impacts on the METC in periods of high and varying congestion and hypercongestion. These results suggest that a single hypercongestion charge is not close to being first best and setting a charge has to take into account the adjustment of demand to hypercongestion.

The issue of the appropriate specification of dependent and independent variables is of even more importance in the case of hypercongestion. If there is bottleneck congestion, the hypercongested relation between speed and flow is represented by a vertical line on the speed-flow diagram. Such bottleneck congestion identified in London implies an elasticity of speed with respect to density of approximately one in magnitude. In the case of error in the measurement of the flow variable, ordinary least square regression of speed against flow will result in a horizontal relation, the opposite of the underlying relation. Thus, it is appropriate to regress the flow variable on speed variable. For the Central London data, there is clear evidence of a bottleneck relation between speed and flow where the flow rate is maintained at the maximum level but speeds vary from very low to high. This raises the question whether in hypercongested urban areas it may be best to analyse queueing at junctions congestion models rather than the standard speed flow model or the model developed in this thesis. This is a question for further research.

By comparison, empirical analysis for the Chalerm Mahanakorn Express Way in Bangkok gives a standard convex hypercongested speed-flow relationship. The hypercongested METC is in the region of 10-100 minutes per vkm. Similar comments apply to those made on the interpretation of the variation in the London estimated METCs.

The final development of density based model is for the marginal external congestion time costs caused by different vehicle types in ordinary congestion. The elasticities of speed with respect to density are still the key parameters in determining the extent of these externalities. The theoretical study showed how the determination of METCs requires consideration of how different vehicles' speeds and densities interact. This involved the use of matrix algebra. Thus, the consequent empirical study has limitations in that speeds are only observed of cars and not the other two vehicle types, though flows are measured for all vehicle types. This is likely to be a problem for all studies as measuring speeds and flows for all vehicle types is a difficult task. In spite of this difficulty, there is a suggestion that the impact on congestion of large and special vehicles may be in excess of the commonly used factor of two or three passenger car equivalents. However, the small number of observations and unavailable observations of speeds for large and exempt vehicle types should be noted.

6.3 Final Conclusions

The ambitious purpose of this thesis has been to develop a sound economic model that provides a basis on which to estimate marginal external congestion time costs. In particular, it is important to develop a model that explains the marginal external congestion time costs of ordinary congestion and hypercongestion and the marginal external congestion time costs of different types of vehicles. In so doing, it is attempted to resolve some of the debates and questions raised in the literature and past studies. The argument of the thesis is that congestion is caused by the density of road occupation and a multiperiod model is required. It is shown that a density based model can be used in empirical investigation. It is shown that, in periods of low congestion with nearly constant speeds and densities, the standard speed flow model and density models predict the same external effect. However, simulation exercises show that the density based model provide a basis for estimating METCs in periods of high and varying congestion and hypercongestion. These estimates are large, vary greatly between the beginning and end of the period of hypercongestion and the assumption about the degree of deterred demand have large effects on the externality. These effects all have implications for the construction of congestion charging schemes and are worthy of further investigation.

The evidence for London suggests bottle neck queuing is the cause of hypercongestion. However, the Bangkok Expressway data suggests that hypercongestion is characterised by a backward bending speed flow curve. It would be useful to consider how appropriate are the speed flow and speed density models in investigating urban traffic networks in cities like London where queuing at intersections is widespread.

The model developed here for determining the METCs of different vehicle types is complex and different from previous methods for investigating such external effects. The data necessary to estimate such METCs is not generally available but the little evidence investigated in this thesis suggest that the passenger car equivalences of larger vehicles may be greater than previously suggested in the literature. These effects are important and are worthy of additional investigation with more extensive data sets.