‘Development (*not under?*) Control’: An econometric model of trip generations at UK office developments.

David C. Broadstock[a], Alan Collins[b*], Lester C. Hunt[a]

JEL Codes: R41 (Transportation demand),

Key words: *Office blocks, missing data, Development Control, Trip generation.*

[a] Surrey Energy Economics Centre (SEEC), Department of Economics, University of Surrey, Guildford, GU2 7XH, United Kingdom.

[b] Department of Economics, Portsmouth Business School, University of Portsmouth, Richmond Building, Portland Street, Portsmouth, Hampshire, PO1 3DE, United Kingdom.

* *Corresponding Author:* Alan Collins, Department of Economics, Portsmouth Business School, University of Portsmouth, Richmond Building, Portland Street, Portsmouth, Hampshire, PO1 3DE, United Kingdom.

‘Development (*not under?*) Control’: An econometric model of trip generations at UK office developments.

Abstract

Development control and wider planning issues are underpinned by a desire to facilitate energy efficient societies and, therefore, incorporating sustainable transport principles directly into development practices. This paper provides compelling evidence of interest to development controllers and academics alike that, at a strategic planning level of observation, there is no evidence that public transport services are provided effectively to new office developments. Rather the results included within this work imply that it is not the carrot (new public services) but rather the stick (reducing parking provision) which elucidates the required sustainability objectives. In arriving at this result, the present paper draws upon the most comprehensive development control database in the UK and also provides commentary on the practical implications of dealing with attrition within a robust and rigorously tested econometric model of vehicle trip attractions.

1. Introduction

The trip to work (and accompanying return trip) continues to serve as an area of hotly contested debate. One perspective might sense some resonance in the work of Shortle and Horan (2002). They define the concept of 'non-point pollution' as pollution that can be specifically attributed to a unique entity, albeit not necessarily generated at the entity's site of residence. In the instance of office blocks this could be described as the pollution generated for the purpose of fulfilling its core business activities but which are not borne at the office block, i.e. the negative externalities caused by their employees travelling to the office, which is a necessary trip to ensure that the office blocks' business activities/requirements are achieved. Hence the pollution generated in the trip to work is a burden which the office block should in some way bear as its own, albeit the decision over which mode employees use to travel to work is often beyond the scope of the employer to control (at least in the short term).

Another key feature of this paper is the formal accounting for land zone placement (i.e. identifying if a site is in the town centre or perhaps a development zone) and which is found to better specify trip generation models.

The organisation of the paper is as follows. The next section defines the general modelling framework. Section 3 provides a description of the dataset with the next section reporting the results. A discussion and concluding remarks are presented in the final section.

2. Modelling Strategy and Estimation

Due to small sample issues and evidence of non-normality in early results, a semi-parametric modelling approach is adopted which is based upon a log-linear equation using a bias-corrected least squares estimator. The semi-parametric approach uses a bootstrap algorithm to generate the full and exact distributions of the econometric model's estimated coefficients.

The General Model

Model I

$$\ln(T_{m,o,d}) = \mu_d + \sum [B^{-1} \sum_{b=1}^{B} (\beta_X^*, \beta_Y^*, \beta_Z^*)]' (\ln X_d, \ln Y_{o,d}, Z_d) \qquad (1)$$

where

- *ln* denotes the natural logarithm of a variable,
- *T* is the total daily traffic flow to a site,
- subscript *m* denotes the mode of travel where $m \in M$ and *M* is the set of travel modes
- subscript *o* denotes the trip origin where $o \in O$ and *O* is the set of origins,
- subscript *d* denotes the destination (i.e. the office development) where $d \in D$ and *D* is the set of destinations,
- μ is the stochastic error term,
- $(\beta_X, \beta_Y, \beta_Z)$ are parameter vectors relating to the vectors of independent variables (X, Y, Z),
- * indicates the sum of the coefficients from each of the *B* bootstrap replications,
- *X* is the Gross Floor Area (GFA) of each site,
- *Y* is a vector of more general site specific characteristics and surrounding area socio-economic characteristics.
- The vector *Z* is formed of 9 dummy variables relating to the land zones identified below, hence no logs are taken.

B = 100, 000 and is the number of bootstrap replications and (*) denotes the sum of the estimated coefficients from each of the bootstrap replications. Each replication creates a new data set with the same dimensions as the original dataset (i.e. the same number of observations and variables). However, each cell is uniquely drawn with replacement from the original dataset, with each individual observation in the original dataset having an equal probability of being drawn. For further explanation of the tenets of bootstrapping processes, see for example Efron and Tibshirani (1993), or for a brief overview of bootstrapping in econometrics see the survey article of MacKinnon (2002).

As this semi-parametric specification produces the exact distribution of the coefficients in the model it thus allows for exact inference to be conducted, rather than approximations based on a normality assumption. This is done using an achieved significance level (ASL), see Efron & Tibshirani (1993), which are analogous to standard p-values for mainstream significance tests, but for the exception that they are exact with respect to the data sample and functional form in question.

This study forms part of a larger research project considering the hypothesis that the demand for travel to a site is influenced by the sites geographical placement, represented by its land zone class. As in Black *et al* (2007a) the hypothesis that $T_{m,o,d} = f(X,Y,Z)$ is applied, coupled with the conjecture that the decision of `where to work' is substantially less elastic than in the case of `where to shop' and may result in land zone placement effects being less prominent for office developments. Gross Floor area for each site is uniquely identified within the model to allow for the specification of an engineering style trip rate regression i.e. $T_{m,o,d} = f(X)$. For completeness the nesting structure where trips are determined purely by site size and location is tested as an alternative i.e. $T_{m,o,d} = f(X,Z)$. This produces three further model specifications;

Model II

$$\ln(T_{m,o,d}) = \mu_d + \sum [B^{-1} \sum_{b=1}^{B} (\beta_X^*, \beta_Z^*)]' (\ln X_d, Z_d) \tag{2}$$

Model III

$$\ln(T_{m,o,d}) = \mu_d + \sum [B^{-1} \sum_{b=1}^{B} (\beta_X^*, \beta_Y^*)]' (\ln X_d, \ln Y_{o,d}) \tag{3}$$

Model IV

$$\ln(T_{m,o,d}) = \mu_d + \sum [B^{-1} \sum_{b=1}^{B} (\beta_X^*)]' (\ln X_d) \tag{4}$$

Three simplifying assumptions are made;

- **A1**: *M* is constrained to passenger vehicle traffic only.
- **A2:** *O* is not known with certainty, and is therefore assumed to be a function of the surrounding areas characteristics.
- **A3:** *D* is constrained to one type of destination, in this instance food superstores. i.e. the model is estimating the levels of trips for only one individual type of activity.

The dependant variable *T* is the total passenger vehicle arrivals (cars) for the period of the survey[1], *Z* is a vector of land zone indicators, *Y* is a vector of more general site specific characteristics and the surrounding areas socio-economic characteristics.

The most general model (Model I, where specifications II, III and IV are nested within it) is conducive with the activity-based theory of travel demand through its constrained analysis of a single activity type. The inclusion of socioeconomic characteristics in the model reflects constraints on time and income within households and furthermore the demand to participate in work, i.e. a larger household has a more extensive and diverse demand requirements than a smaller household thereby requiring a larger income for their general maintenance. It is thus conjectured that the demand to travel to a given activity by a specific mode of transport (to *work* by *car*) is therefore likely to be subject to a set of estimated parameters which are heterogeneous to alternative activity purposes.

The semi-parametric approach to estimation allows for far more detail to be derived on the distributions of the model coefficients, which means that the assumption of a normal distribution can be relaxed. In short the approach to inference derives the coefficient distribution for the specific data at hand given the estimation method applied.

3. Data

[1] The hourly average is not considered as this would merely rescale the dependent variable, as all surveys are of the same length.

The data used to analyse this problem is a composite dataset assembled to investigate the determinants of car borne UK office development trip rates. Transport analysts and planners on the ground have to make rational defensible cases for decisions on the acceptance/rejection or scale of land use developments at particular sites. Typically they rely on comparators drawn from 'similar sites' elsewhere. In an attempt to systematise and expand the range of available comparator sites, the Trip Rate Information Computer System (TRICS) has evolved. This is essentially a UK based mirror (though with more detailed site specific information) of the US Institute of Transport Engineers (ITE) trip rate generation database. Another element of the composite dataset comprises socio-economic data from publicly available sources. As such it represents the data that in principle could be readily used by the most informed transport planners/engineers. In short, it represents the best available data without commissioning new and costly travel/traffic surveys. Nevertheless, this data features a considerable number of missing elements.

The composite dataset therefore provides a cross-section of information for 50 office block developments over the period 1987-2002. Data is taken from the TRICS database and combined with information from the UK Census at the Office of National Statistics and NOMIS (National Online Manpower Information System). Descriptive statistics of the full dataset are provided in Table 1. This table exemplifies that missingness is persistent in the dependant variable and that 15 out of a potential 50 observations are missing. This is not uncommon in such data and the empirical estimations apply a range of methods are presented in the literature to overcome this problem. In order that the subsequent analysis holds as much value as possible eight of these methods (also known as imputation methods) are applied in the present example and are summarised in Black *et al* 2007a. The subsequent analysis of the results will also make reference to the impacts of the different imputation methods.

[INSERT TABLE 1 ABOUT HERE]

Figure 1 provides the mean average of the total trip arrivals and departures of the course of the

24 hour period and indicates that the sample, although relatively small, reflects a-priori expectations of vehicular movement. The data reflects that the majority of surveys are conducted around the `normal' working hours periods (i.e. based around 9am-5pm), however a number of the sites are known to be in operation 24 hours a day resulting in arrivals exceeding departures marginally. Figure 1 reveals evidence of bi-modal peaks for both the arrivals and departures, which are situated around the start and end of the work and lunch periods. The start of work peak in average arrivals occurring at 09:00, though spread from 06:30-10:00. Similarly the end of work peak in average departures is at 17:30, with an observably more concentrated spread of trip generations starting from around 16:30 and finishing at 18:30. The lunchtime network traffic loading occurs between 12:00 and 14:00, with the peak sizes suggesting that there is a significant proportion of employees who do not return to work after lunch. Reasons for this may include the end of a shift, the need to visit another office in the afternoon and/or the decision to take a half day or continue working the rest of the day from home (teleworking). The implications of these different reasons in terms of their contribution to non-point pollution and negative externalities are diverse. For instance, taking a half day and/or teleworking for half a day still necessitates two journeys (from home to work and the associated return journey). In the instance of a half day, this means almost the same amount of non-point pollution is generated as per the employees working whole days, with only half the on-site productivity. The slight reduction in externalities will be a result of one end of the return journey being made during the off peak period, thus providing less network friction (congestion etc.).

[INSERT FIGURE 1 ABOUT HERE]

4. Results

This section proceeds by presenting the results for model specifications I, II, III and IV sequentially[2]. Each model specification is initially analysed so as to observe the effects of the imputation methods (denoted (a)-(h)) on the qualitative results, looking specifically for changes in the magnitude and/or

[2] Note that for the ASL's, the unimputed (or listwise cleansed data) uses n=35, while for the imputed data n=50 for determination of the critical value of the test statistic ($t(z^*)$).

sign of the estimated elasticities. The coefficients reported in each of these tables are the bias corrected least squares estimates.

[INSERT TABLE 2 ABOUT HERE]

[INSERT TABLE 3 ABOUT HERE]

[INSERT TABLE 4 ABOUT HERE]

[INSERT TABLE 5 ABOUT HERE]

Table 2 provides the estimates the most general model considered in this study, implying that trip rates are a function of site characteristics, socio-economic and demographic information and also spatial placement. The results are not particularly robust across the alternative imputation methods used, with the majority of variables producing elasticities that differ not only in magnitude, but also in sign. For instance, the elasticity estimates upon gross floor area lie within the range [-0.427, 0.689], although it must be recognised that none of the estimates satisfy standard hypothesis tests.

Model cross-comparison is done to infer which model (I,II, III or IV) best explains vehicle trip rates given the data available? The criteria applied here namely the Root Mean Squared Error (*RMSE*), adjusted R-squared ($\overline{R}^2$), the Akaike information criterion and Schwartz-Bayes information criterion tests (*AIC* and *BIC)*, for further explanation on these see for example Greene (2003) or Diebold (2004). No natural testing process is known to establish which method, (a)-(h), is superior,[3] however these results are presented for completeness and further to highlight the impacts of analyst choices. Hence, the emergent story from the present application is that no model can be

[3] In situations where substantial quantities of data are available an amount can be withheld for predictive comparisons, however given the small size of the present data this was not considered feasible.

considered *strictly* preferred based on the results presented in this paper. Given the four model selection criteria considered, model I is the *most* preferred specification given the data available.

A sensitivity analysis is conducted using the results arising from Specifications III and IV[4], illustrated in Figure 2. The two most extreme estimates, given the alternative imputation methods, are found in these specifications, i.e. the two (statistically significant) estimates of δ with the greatest distance between them. This figure indicates the range of predicted trip rates when GFA is adjusted from the minimum value observed in the dataset, through to the maximum using a consistent estimation and inference methodology, but conflicting imputation methodologies. It thus exemplifies the situation where, what may at first appear as an almost negligible difference in parameter estimates, can in fact have more severe consequences when applied. The associated parameters are identified in the figure and the *ceteris paribus* impact upon trip rates at the uppermost extreme point of observed site sizes are as follows for the specifications considered; (i) III(a)=4.60, (ii) III(f)=5492.10 [specification III range^{max}=5487.50 trips, indicated by dashed region], (iii) IV(d)=2076.24, (iv) IV(e)=1081.53 [specification I range^{max}=994.71 trips, indicated by the dotted region].

[INSERT FIGURE 2 ABOUT HERE]

5. Discussion and Concluding Remarks.

This paper has presented an analysis of vehicle trip attractions at a representative sample of UK Office Block developments in the UK. In so doing it has provided a framework to elucidate the impact of public service provision at the chosen sites, i.e. at a coarse, or rather from a different perspective ‘strategic’, level view of single developments, there is no evidence to reinforce the proposition that public service provision is anything other than a blunt instrument for facilitating sustainable travel. Through a range of alternative model specifications, and applying a range of methods to expand the dataset as far as possible, the evidence suggest at best that providing

[4] These specifications are chosen as they provide consistently significant representations of the variable in question.

public services has a zero effect on trip generations by car, and at worst it is associated with increases in car use. From a more optimistic perspective there is thus arguably evidence that use of public transport could still be better fostered in order to underpin sustainability objectives by implementing a suite of policies and toolkits which are not simply blunt instruments implemented to supported alternative agendas.

Assuming consumers are rational agents and that they are maximizing their utility over time and monetary budget constraints, the underlying conclusions in this paper are that to date, initiatives to encourage the use of public transport to travel to work have not proved effective. The availability of bus services at the work-place is not being met with decreased use of the car. The implications of this are that the monetary price of the bus services is not low enough to attract a decent quantity of demand, and/or the supply chain for this mode of transport does not match the consumers requirements.

Acknowledgements

Support from the UK Economic and Social Research Council (Project Number PTA-033-2004-00035) is gratefully acknowledged, as is permission from JMP Consultants Limited and the UK TRICS Consortium, to use the TRICS database.

References

Allison, P. (2000), Multiple imputation for missing data, *Working paper: University of Pennsylvania.*

Black, C., Broadstock, D. C., Collins, A. & Hunt, L. C. (2007*a*), 'Filling in the gaps' in transport studies: A practical guide to developments in data imputation methods, *Traffic Engineering and Control* **48**(7), 64-68.

Black, C., Broadstock, D. C., Collins, A. and Hunt, L. C. (2007*b*), The derived demand for traffic at food superstores in the UK: A semi-parametric regression approach, *International Journal of Transport Economics **34***(3), 403-427.

Brownstone, D. (1998), Multiple imputation methodology for missing data, non-random response, and panel attrition, pp. 421-450 in Garling, T., Laitila,.T and Westin, K., (1998)*, Theoretical Foundations of Travel Choice Modelling*, Elsevier, UK, .

Davidson, R. and Mackinnon, J. (2004), *Econometric theory and methods*, Oxford University Press, New York.

Davison, A. and Hinkley, D. (1997), *Bootstrap Methods and Their Applications*, Cambridge University Press.

Diebold, F. (2004), *Elements of Forecasting*, Thomson South Western, Canada.

Efron, B. and Tibshirani, R. (1993), *Monographs on Statistics and Applied Probability 57: An Introduction to the Bootstrap*, Chapman and Hall/CRC, USA.

Greene, W. (2003), *Econometric Analysis*, 5th edition, Prentice Hall, USA.

Gujarati, D. (2003), *Basic Econometrics*, 4th edition, McGraw Hill, London.

Kofman, P. & Sharpe, I. (2000), Imputation methods for incomplete dependent variables in finance, *Working paper: University of Technology, Sydney*.

Lavori, P., Dawson, R. and Shera, D. (1995), `A multiple imputation strategy for clinical trials with truncation of patient data', *Statistics in Medicine* **14**, 1913-1925.

Little, R. and Rubin, D. (2002), *Statistical Analysis with Missing Data*, John Wiley and Sons Inc., USA.

MacKinnon, J. (2002), Bootstrap inference in econometrics, *Canadian Journal of Economics* **70**.

Rosenbaum, P. and Rubin, D. (1983), The central role of the propensity score in observational studies for causal effects, *Biometrika* **70**, 41-55.

Rubin, D. (1981), The bayesian bootstrap, *The Annals of Statistics* **9**(1), 130{134.

Rubin, D. and Schenker, N. (1986), Multiple imputation for interval estimation from simple random samples with ignorable nonresponse, *Journal of the American Statistical Association* **81**(394), 366-374.

Shortle, J. S. and Horan, R. D. (2002), The Economics of Nonpoint Pollution Control, pp. 4-40, in Hanley, N. and Roberts, C., *Issues In Environmental Economics*, Blackwell Publishing, UK,

Wood, A., White, I., Hillsdon, M. and Carpenter, J. (2005), `Comparison of imputation and modelling methods in the analysis of a physical activity trial with missing outcomes', *International Journal of Epidemioplogy* **34**, 89-99.

Appendix A - Land Use Definitions

The following information is reproduced with permission from the TRICS consortium, and is taken from the 2006(a) version of the database.

[INSERT FIGURE 3 ABOUT HERE]

TRICS LOCATION DEFINITIONS - JULY 2005

Town Centre Within the central core area of the heart of the town e.g. the primary shopping area, as defined in the local development plan.

Edge of Town Centre For retail, a location within easy walking distance (i.e. up to 300 metres) of the primary shopping area, often providing parking facilities that serve the centre as well as the site, thus enabling one trip to serve several purposes. For other uses, edge-of-centre may be more extensive, but within 300m of the town centre boundary, based on how far people would be prepared to walk. For offices this may be outside the town centre but in the urban area within 500m of a public transport interchange. Local topography and barriers will affect pedestrians' perceptions of easy walking distance. Examples of barriers include crossing major roads and car parks. The perceived safety of the route and strength of the attraction of the town centre are also relevant.

Neighbourhood Centre (Local Centre) Residential area, similar to "Suburban Area", but with additional amenities like local shops, schools, etc. Could be described as a small "district" or "village" within the town itself. Would also apply to actual villages. The Local centres include a range of small shops of a local nature, serving a small catchment. These may include a general grocery store, a newsagent, a sub-post office and a pharmacy. These centres provide accessible shopping for people's day-to-day needs.

Suburban Area (Out of Centre) A Residential area that is outside the town centre, but not at the physical edge of the town itself. Villages are included as Neighbourhood Centre. Edge of Town At the physical edge of the town/city, where the town meets the countryside.

Free Standing (Out of Town) Out of town, beyond the physical edge of the nearest town/city, in the countryside.

Commercial Zone An area of significant business activity within a town.

Industrial Zone An area of significant industrial activity within a town.

Development Zone An area of redevelopment or regeneration, for example London Docklands (or on a smaller scale for other towns and cities).

Tables

Table 1: Descriptive Statistics for Unimputed Data

Continuous Variables	Mean	Median	s.d.	Min	Max	N
$\ln Arr$	5.550	5.727	1.055	2.734	6.993	35
$\ln GFA$	8.638	8.627	1.087	6.011	11.054	50
$\ln SEMP$	5.746	5.851	1.050	3.611	7.937	50
$\ln PARK$	5.135	5.244	1.220	2.303	7.601	50
$\ln CAR$	-0.176	-0.223	0.372	-1.386	0.262	50
$\ln PSV$	3.738	4.248	1.214	0	4.248	50
$\ln AVHS$	0.904	0.904	0.055	0.792	0.988	50
$\ln AEMP$	11.113	10.998	0.603	9.977	12.282	50
$\ln APOP$	11.927	11.800	0.578	10.999	12.987	50
$\ln HHEMP$	0.091	0.088	0.130	-0.154	0.431	50
$YEAR$	1993.700	1992	4.617	1987	2002	50

$(\ln T_{m,o,d}) = \ln ARR$, $(\ln \delta_d) = \; lnGFA$
$(\ln \gamma_d)$=*all other continuous variables except year which was consistently insignificant and hence omitted from the models*

Discrete Variables	Mean	Median	s.d.	Min	Max	N
$LU1$(base)	.18	0	0.388	0	1	50
$LU2$	.04	0	0.198	0	1	50
$LU3$	.12	0	0.328	0	1	50
$LU4$	.26	0	0.443	0	1	50
$LU5$	.1	0	0.303	0	1	50
$LU6$	.08	0	0.274	0	1	50
$LU7$	.06	0	0.240	0	1	50
$LU8$	.06	0	0.240	0	1	50
$LU9$	.1	0	0.303	0	1	50

$(\ln \lambda_d) = (LU1, ..., LU9)$
These classifications are consistent with government planning policy guidance documents, further explanation given in Appendix A.

Table 2: Spec I results: $T_{m,o,d} = f(\delta, \gamma, \lambda)$

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
lnGFA	-0.346	0.027	-0.427	0.689	-0.233	0.230	-0.308*	0.251
lnSEMP	0.381	0.112	0.383	-0.107	0.260	-0.015	0.350***	0.047
lnPARK	0.898*	0.471	0.623*	0.104	0.628	0.734**	0.898***	0.291
lnCAR	-0.057	-0.493*	-0.698*	-1.226*	-0.719	0.120	-0.046	-0.887**
lnPSV	0.188	-0.025	0.060	0.017	0.357	0.128	0.127*	0.054
lnAVHS	2.760	4.690*	2.208	6.166	1.777	4.356*	2.820**	5.181*
lnAEMP	-0.134	-0.215	-0.080	-0.479	-0.322	-0.291	-0.144*	-0.276
lnHHEMP	-0.541	0.210	0.317	1.120	0.491	0.345	-0.348	0.594
LU1	-0.299	0.192	2.009	-0.671	2.824	-1.190	-0.131	-0.611
LU2	0.633	0.799	1.494	0.265	2.494	-0.429	1.053	0.212
LU3	0.177	0.810	2.820	-0.364	3.716	-1.142	0.335	0.116
LU4	-0.094	0.910	2.764	0.329	3.188	-1.178	0.091	0.260
LU5	-0.119	0.245	2.394	-0.532	2.813	-1.465	0.036	-0.443
LU6	0.056	-0.065	2.212	-1.353	3.514	-1.517	0.081	-0.678
LU7	0.125	0.228	1.870	-0.487	3.031	-1.130	0.306	-0.478
LU8	0.510	-0.212	2.143	-1.374	4.113	-1.162	0.408	-0.703
LU9	0.456	0.112	2.179	-1.683	3.031	-1.212	0.580	-0.963
Obs	35	50	50	50	50	50	50	50
$RMSE$	0.21	0.41	0.42	0.58	0.69	0.33	0.17	0.45
$\bar{R}^2$	0.92	0.68	0.712	0.58	0.36	0.87	0.96	0.63
AIC	0.12	0.33	0.343	0.66	0.95	0.22	0.06	0.40
BIC	0.25	0.62	0.656	1.27	1.81	0.42	0.11	0.77

Notes;***=1%,**=5%,*=10%

Table 3: Spec II results: $T_{m,o,d} = f(\delta, \gamma)$

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
lnGFA	-0.198	-0.323	n/a	-0.248	-0.289	-0.250	-0.250	-0.396
lnSEMP	0.321	0.457**	n/a	0.336	0.372	0.349**	0.349**	0.343
lnPARK	0.703***	0.498***	n/a	0.580*	0.571**	0.722***	0.722***	0.682***
lnCAR	0.092	-0.438	n/a	-1.124**	-0.482	0.040	0.040	-1.031**
lnPSV	0.063	0.096*	n/a	0.245***	0.063	0.065**	0.065**	0.117**
lnAVHS	2.372*	3.862***	n/a	4.177*	2.466	2.325***	2.325***	3.173**
lnAEMP	-0.044	-0.055	n/a	-0.144	0.036	-0.024	-0.024	0.013
lnHHEMP	-0.556	-1.456*	n/a	-0.873	-0.673	-0.505	-0.507	-0.756
Obs	35	50		50	50	50	50	50
$RMSE$	0.30	0.44		0.71	0.719	0.25	0.25	0.49
$\bar{R}^2$	0.90	0.72		0.53	0.45	0.93	0.93	0.69
AIC	0.14	0.26		0.67	0.71	0.09	0.09	0.33
BIC	0.20	0.36		0.90	0.97	0.12	0.12	0.45

Notes;***=1%,**=5%,*=10%

Table 4: Spec III results: $T_{m,o,d} = f(\delta, \lambda)$

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
lnGFA	0.138***	0.594***	0.444***	0.532***	0.497***	0.779***	0.258	0.577***
LU1	1.202	0.194	1.164	1.266	0.761	-1.567***	3.046**	0.112
LU2	0.738	0.631	0.924	1.652	0.739	-0.538	2.775**	1.231
LU3	1.298	0.750	2.122**	1.315	1.643	-0.936*	3.824**	0.910
LU4	1.303	0.813	2.462***	2.038	1.488	-0.880*	3.861**	0.728
LU5	1.230	0.405	1.870*	1.499	1.139	-0.886*	2.998*	0.318
LU6	1.322	0.328	1.958*	1.183	1.463	-1.110*	3.067*	0.428
LU7	0.976	0.417	1.275	1.232	1.120	-0.774*	2.577*	0.365
LU8	1.115	0.382	1.521*	0.855	1.125	-1.009**	2.928*	0.519
LU9	1.158	-0.370	0.831	0.126	0.414	-1.841***	2.316	-0.229
Obs	35	50	50	50	50	50	50	50
$RMSE$	0.44	0.49	0.50	0.71	0.74	0.37	0.77	0.61
$\bar{R}^2$	0.76	0.63	0.68	0.48	0.39	0.83	0.24	0.48
AIC	0.34	0.36	0.37	0.75	0.82	0.20	0.89	0.56
BIC	0.53	0.52	0.55	1.10	1.21	0.29	1.31	0.82

Notes;***=1%,**=5%,*=10%

Table 5: Spec IV results: $T_{m,o,d} = f(\delta)$

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
lnGFA	0.648***	0.637***	n/a	0.691***	0.632***	0.648***	0.648***	0.642***
Obs	35	50		50	50	50	50	50
$RMSE$	0.60	0.72		0.90	0.83	0.50	0.50	0.53
$\bar{R}^2$	0.66	0.39		0.32	0.37	0.72	0.72	0.71
AIC	0.38	0.55		0.85	0.72	0.26	0.26	0.29
BIC	0.4	0.57		0.88	0.75	0.27	0.27	0.30
Notes;***=1%,**=5%,*=10%								

Figure Legend

Figure 1: Total car traffic arrivals by time of day
[Insert file “offarrdep.eps”]
Figure 2: Sensitivity of trip rate estimates to changes in imputation method used
[Insert file “sensitivity.eps”]
Figure 3: TRICS location definitions
[Insert file “Tricsludef.eps”]

Figure 1

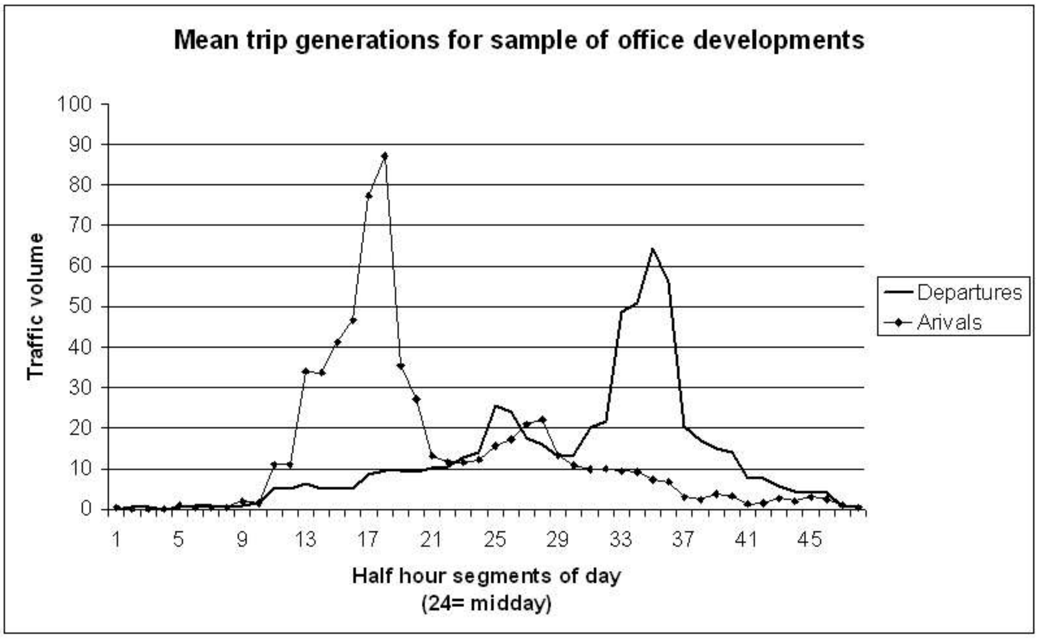

Mean trip generations for sample of office developments
100
90
80
70
60
50
40
30
20
10
0
Traffic volume
1
5
9
13
17
21
25
29
33
37
41
45
Half hour segments of day
(24= midday)
Departures
Arivals

Figure 3

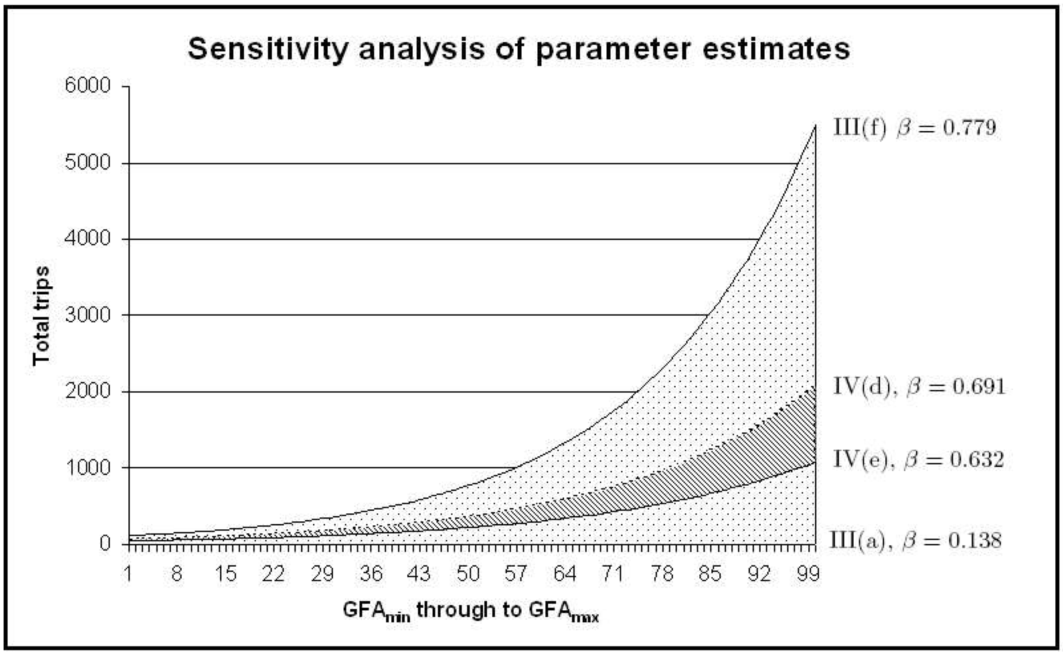
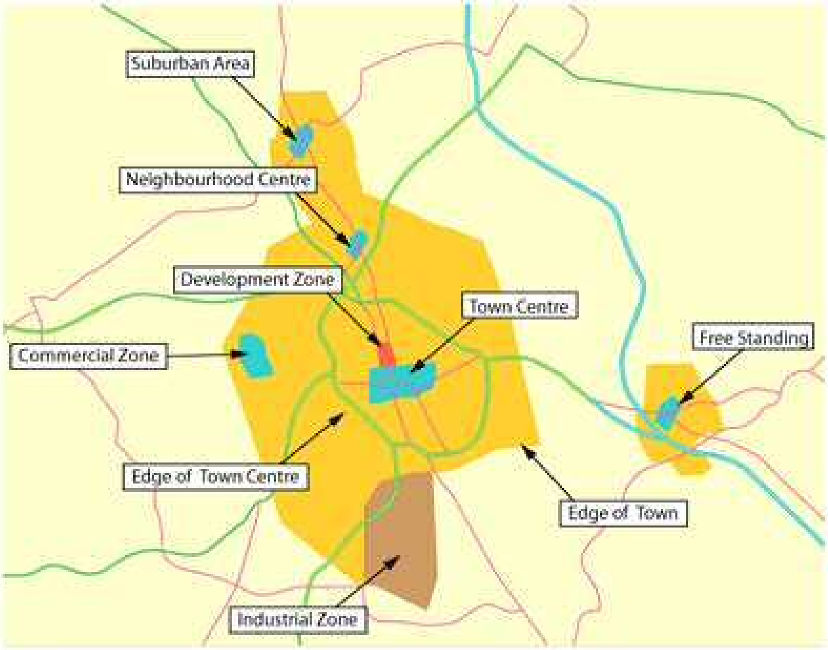

Sensitivity analysis of parameter estimates
Total trips
6000
5000
4000
3000
2000
1000
0
1 8 15 22 29 36 43 50 57 64 71 78 85 92 99
GFAmin through to GFAmax
III(f) β = 0.779
IV(d), β = 0.691
IV(e), β = 0.632
III(a), β = 0.138

Figure 3



Suburban Area
Neighbourhood Centre
Development Zone
Town Centre
Free Standing
Commercial Zone
Edge of Town Centre
Edge of Town
Industrial Zone

David

Thanks for producing the revised paper that I think is starting to look a lot better than the submission to TRA; however, I worry that the key messages are not clear. My reading is as follows:

- We are aim trip attraction models for UK office developments in order to ascertain what the key 'drivers' are for the trips. Within this, we are able to test the assertion that 'bus services need to be provided at newly opened sites'.

- To do this we set up a trip attraction model. To estimate we use a unique database (TRICS + ONS). However, for statistical reasons we use bootstrapping. But there is also a problem of missing data, so we use 'missing data' techniques to solve (and check the robustness of the results).

- We find, in addition to the 'drivers' of trips that the 'bus services' argument is wrong.

Is this it? Or have I got it wrong? If I have not then it does not come through as clear as maybe it could (but I may well have misunderstood it).

I think we need to simplify the message somewhat and stick to finding out about the economics, rather than being bogged down in techniques. I therefore suggest the following/or have the following queries:

- **Abstract**: Could it be more precise/to the point. That said there is no mention of empirical testing or econometrics.

- **Title**: Not sure I understand the question mark etc.

- **Introduction**: Is there too much on data (second para)? Shouldn't it concentrate on the economic issues?

  Take out reference to Says Law

- **Section 2 Data**: This comes almost straight after intro para on data – to me it does not work. Furthermore, I think Section 3 Modelling should come before the data.

  In this section, however, can't we just state the problem of the missing data and say there are n techniques to solve this (referencing the previous literature and our TEC paper (Black et al 2007b)? I.e. we state that there are the n techniques and we use x of them in order to check the robustness of the results. And then just eliminate the material in the appendix pages 16-18? Or am I missing something?

- **Section 3 Modelling etc.**: I strongly think this should come before the data section – otherwise it looks if we are just driven by data.

Page 6 we have μ_d and $_\mu$ is this right?

Page 6 (and onwards) we still have Greek letters for variables. Suggest this is changed to X_1, X_2, etc.

Page 7. Should the statement about part of wider research have come in the introduction?

Pages 19-20 in appendix. Are these needed? Couldn't we just refer to previous work? Or again am I missing something?

- **Section 4 Results**: Page 9, personally do not like the term "These are well expounded in the literature", suggest we just refer to the texts .

  Page 10/last para of section. Could this be clearer? Do we need a summary of the results before going to the conclusion?

- **Section 5. Discussion and conclusion**: Is this enough? Isn't there more to say on the economics? Does the conclusion about buses come out enough?

Sorry David that this sounds a bit negative. I think there is the basis for a nice paper, but we need to get the message clearer. What do you think Alan – am I being over fussy?

Cheers

Lester