White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/2155/

Monograph:
Bristow, A.L., Fowkes, A.S., Bonsall, P.W. et al. (1994) The Optimisation of Integrated Urban Transport Strategies: Tests Using Pluto. Working Paper. Institute of Transport Studies, University of Leeds , Leeds, UK.

Working Paper 424

Reuse
See Attached

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

White Rose Research Online
http://eprints.whiterose.ac.uk/

Institute of Transport Studies
University of Leeds

This is an ITS Working Paper produced and published by the University of Leeds. ITS Working Papers are intended to provide information and encourage discussion on a topic in advance of formal publication. They represent only the views of the authors, and do not necessarily reflect the views or approval of the sponsors.

White Rose Repository URL for this paper:
http://eprints.whiterose.ac.uk/2155/

Published paper

Bristow, A.L, Fowkes, A.S, Bonsall, P.W., May, A.D. (1994) *The Optimisation of Integrated Urban Transport Strategies: Tests Using Pluto.* Institute of Transport Studies, University of Leeds. Working Paper 424

White Rose Consortium ePrints Repository
eprints@whiterose.ac.uk

UNIVERSITY OF LEEDS
Institute for Transport Studies

ITS Working Paper 424 *ISSN 0142-8942*

October 1994

THE OPTIMISATION OF INTEGRATED URBAN TRANSPORT STRATEGIES: TESTS USING PLUTO

A L Bristow
A S Fowkes
P W Bonsall
A D May

ITS Working Papers are intended to provide information and encourage discussion on a topic in advance of formal publication. They represent only the views of the authors, and do not necessarily reflect the views or approval of the sponsors.

CONTENTS

		Page
	ABSTRACT	
1.	INTRODUCTION	1
	1.1 Objectives	1
	1.2 PLUTO	1
2.	ROAD INVESTMENT PROJECTS	3
	2.1 A systematic approach	3
	2.2 Intuitive approach	8
	2.3 Non-modelling approach	10
	2.4 Model performance	10
	2.5 Predicting the optimum	12
	2.6 Conclusions on infrastructure projects	12
3.	CORDON PRICING	13
	3.1 General data set	13
	3.2 Systematic approach	14
	3.3 Conclusions on cordon pricing	17
4.	BUS FARES AND SERVICE LEVELS	18
	4.1 Models tested	18
	4.2 Conclusions on bus fares and service levels	19
5.	CORDON PRICING, BUS FARES AND SERVICE LEVELS	21
	5.1 Initial tests	21
	5.2 Cordon pricing, bus fares and bus service levels - interim conclusions	26
	5.3 Revised Tests	27
	5.4 Conclusions on bus fare, service level and cordon pricing tests	31
6.	INTRODUCING FURTHER COMPLICATIONS	32
	6.1 An additional independent variable	32
	6.2 Behaviour of other dependent variables	33
7.	CONCLUSIONS	36
	7.1 Discrete policy variables	36
	7.2 Continuous policy variables	36
	7.3 Multiple policy variables	36
	7.4 The optimisation process	36
	7.5 General conclusions	38
	REFERENCES	39
	APPENDIX 1 - ROAD PROJECT MODEL PARAMETERS	40
	APPENDIX 2 - CORDON PRICING	41
	APPENDIX 3 - FORMAL RULE FOR TESTING	43
	APPENDIX 4: BUS FARES AND SERVICE LEVELS	45
	APPENDIX 5: COMBINED TESTS - OPTIMUM INPUT LEVELS	46

ABSTRACT

BRISTOW, AL, FOWKES, AS, BONSALL, PW, MAY, AD (1994). The optimisation of integrated urban transport strategies: tests using PLUTO. *ITS Working Paper 424,* Institute for Transport Studies, University of Leeds, Leeds.

This working paper reports work indertaken on an EPSRC study, the optimisation of integrated urban transport strategies. It is the first in a series of papers each reporting work using a different transport model. This paper is concerned with work using PLUTO, a transport model based on a hypothetical city, which can provide model runs quickly and cheaply.

PLUTO was used to experiment widely, examining paths towards an optimum when considering discrete policy variables, continuous policy variables and hybrids where several policy variables may be combined to form a strategy.

Our search for a method by which to reach an optimum solution uses regression analysis of carefully specified sets of model runs. We find that the use of statistical modelling techniques is extremely useful in pointing the way to an optimum, using only a limited number of model runs. However, care is necessary to ensure that the regression models are interpreted correctly.

THE OPTIMISATION OF INTEGRATED URBAN TRANSPORT STRATEGIES: TESTS USING PLUTO

1. INTRODUCTION

1.1 Objectives

This working paper describes work based on the PLUTO urban transport model (Bonsall; 1992; 1994), carried out during an EPSRC study on the optimisation of integrated urban transport strategies. An objective of this study was to identify methods by which an 'optimum' solution (however defined) could be arrived at using as few runs of the transport model as possible.

The modelling work described in this paper has two objectives

(i) to produce a regression model to predict the optimum run within PLUTO given a limited range of policy options
(ii) to identify a process by which the runs necessary to produce a predictive regression model to identify the optimum, may be specified.

PLUTO was an extremely useful tool for the main developmental stage of our study because of the rapidity with which new runs may be completed. A run of PLUTO including setting input values, running the model for each of ten years, running the evaluation package and recording results, takes approximately 20 to 30 minutes. A similar process using one of the strategic transport models available to us occupies approximately three hours. PLUTO was therefore used to develop our approach, which was then tested on the 'real' strategic transport models of Edinburgh and London. The results of the testing process using the Edinburgh model are reported in Bristow et al, 1994.

The variable which we seek to optimise is Net Present Value (NPV), being the most comprehensive single variable available in that it includes those costs and benefits which can be valued. The behaviour of other variables of interest, eg the number of personal injury accidents and environmental impacts is discussed in relation to NPV.

This report is divided into seven sections, each reporting a specific phase of the work. Section 2 examines the issue of discrete infrastructure projects, and describes a process by which the optimum combination may be identified. Section 3 examines the question of cordon pricing, where we are dealing with a continuous variable in the cordon charge and a discrete variable in the choice of cordon boundary from six possible options. Section 4 examines bus fares and service levels, looking at the peak and the off-peak. Section 5 examines a complex problem where cordon price and boundary, bus fares and bus service levels are all allowed to vary. In section 6 we consider adding a further complication in the form of infrastructure investment, and the behaviour of other dependent variables. General conclusions are drawn in section 7.

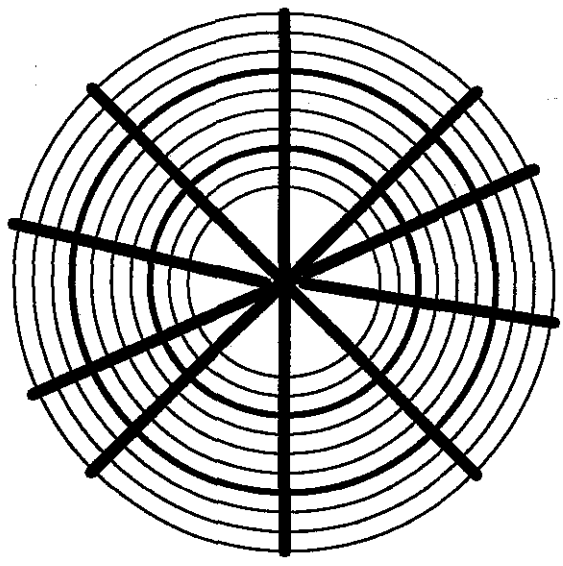
1.2 PLUTO

The PLUTO urban transport model is based on a hypothetical and symmetrical city 'Plutopia', Figure 1.1 shows the main road network and concentric zones. PLUTO was originally designed as an educational tool, running for one year at a time, allowing those using it to make annual policy adjustments. It contains a wide range of policy options, road infrastructure projects,

parking capacity and charging, bus service levels and fares, bus lanes, traffic calming, road safety education, cordon charging, etc.

One of our first requirements was to create a batch version of PLUTO; this runs for 10 years to show the long term policy impacts. It was also necessary to devise an evaluation package to calculate the Net Present Value of any policy combination tried. Once these tasks were completed we were able to begin our transport strategy tests.

Figure 1.1: A Sketch Map of the City of "Plutopia"



Link	Capacity (pcu per hour)	Speed (Average peak period car speed, year 10 do-minimum)
1	2500	7
2	2300	6
3	2200	6
4	2200	4
5	2500	8
6	2500	9
7	2900	13
8	3000	15
9	2400	22
Inner ring road	2400	22
Outer ring road	2600	14

2. Road Investment Projects

PLUTO was set up to allow investment in up to ten road infrastructure projects detailed in Table 2.1. Each scheme is a discrete entity, it is either implemented or not, there are no continuous variables, and a priori we do not know how the projects will interact with each other, except that, as projects 4 and 5 both affect link 4 and projects 8 and 10 both affect the inner ring road, they will to a large extent be alternatives. We have followed two main approaches to identify the PLUTO model runs that will produce the best regression model; a systematic approach outlined in section 2.1 and an intuitive approach discussed in section 2.2.

Table 2.1: Road investment projects available

Project	Cost (£m)	Effect
1	20	Increase link 1 capacity to 3000 pcu/hr
2	15	Increase link 2 capacity to 2800 pcu/hr
3	15	Increase link 3 capacity to 3000 pcu/hr
4	10	Increase link 4 capacity to 2400 pcu/hr
5	16	Increase link 4 capacity to 3200 pcu/hr
6	8	Increase link 5 capacity to 3000 pcu/hr
7	2	Increase link 6 capacity to 3000 pcu/hr
8	15	Increase inner ring road capacity to 3000 pcu/hr
9	12	Increase capacity of outer ring road to 3000 pcu/hr
10	22	Increase capacity of inner ring road interchanges

2.1 A Systematic Approach

This approach begins as follows:

(i) Each of the 10 projects should be tested alone, to establish the NPV of the scheme with no interactions (see Table 2.2).

(ii) One additional run should be conducted containing all 10 schemes, to give an extreme example of interaction (see Table 2.2).

Table 2.2: NPV from PLUTO runs of single projects and all ten together

Project	NPV (£M)
P1	-19.2
P2	-10.3
P3	-19.1
P4	-9.1
P5	+16.4
P6	-0.3
P7	-15.6
P8	+4.3
P9	+88.6
P10	+37.3
P1 to P10	+66.8

Four of the ten available individual schemes have a positive NPV when implemented in isolation. A simple summation of all ten NPVs is £73M, which is higher than the actual NPV of £66.8M when all ten are included in a PLUTO run. This is partly due to the presence of alternatives, for example, if the obvious alternatives P4 and P8, are excluded the sum is £77.8M, while the NPV obtained from a PLUTO run including the remaining eight projects is £96.9M. In this case the whole is greater than the sum of the parts, suggesting that there are significant synergetic interactions between schemes. We do not rule out any of the ten projects at this stage, as those with a poor performance in terms of NPV in isolation may improve in combination.

(iii) Try every possible pair of projects; for 10 projects this meant 45 combinations. This gives the regression model a comprehensive set of data on how any project will interact with any other, all other things being equal.

Once these pairs have been run the PLUTO output can be used to investigate the interaction between pairs of projects. In the search for an optimum we are looking for synergy between schemes. Therefore, at this stage we are seeking to exclude substitutes and include combinations that act as complements. We therefore examined the NPV for each paired test and compared it to that of the better of the two individual projects concerned. For example, suppose Project X alone has an NPV of 10, while project Y alone has an NPV of -5 and if projects X and Y are both implemented the NPV is 12. In this case the incremental value obtained from X + Y as opposed to X alone is 2. In this way we can identify projects which perform poorly alone, such as Y, yet when combined with another project have a positive impact. Similarly, if the joint NPV turns out to be less than one of the individual NPVs we then know that these two projects are acting as substitutes.

Table 2.3, overleaf, shows the results obtained by examining each pair of projects in this way. The NPV outcome from a PLUTO run of each pair is then compared with the NPV from a PLUTO run of the better individual project. The final column in the table gives the difference between the two NPVs. Thus, a negative change indicates an incremental disbenefit from the pairing, while a positive change (marked with a *) indicates the positive synergy which we seek. Table 2.3 (overleaf), shows that there are only twelve positive interactions between projects when tested in pairs.

In Table 2.4, we examine the twelve positive pairs in more detail.

Table 2.3: Project Pairs - Impacts of Interaction on NPV

Project Pairs	Paired test NPV	Best Single Project NPV	ΔNPV	
1 & 2	-15.27	-10.32	-4.93	
3	-15.38	-19.11	-3.73	
4	-17.22	-9.06	-8.22	
5	8.25	16.44	-8.19	
6	-16.27	-0.35	-14.93	
7	-24.86	-15.66	-9.22	
8	-11.88	4.31	-16.19	
9	76.50	88.61	-12.11	
10	21.81	37.31	-15.5	
2 & 3	-25.16	-10.32	-15.39	
4	-15.16	-9.06	-6.1	
5	10.39	16.44	-6.05	
6	-8.13	-0.35	-7.78	
7	-25.71	-10.32	-15.39	
8	-3.74	4.31	-8.05	
9	80.03	88.61	-8.58	
10	27.42	37.31	-9.89	
3 & 4	-24.73	-9.06	-15.67	
5	13.85	16.44	-2.59	
6	-6.35	-0.35	-6.0	
7	-16.25	-15.66	-0.59	
8	-11.84	4.31	-16.15	
9	79.97	88.61	-8.64	
10	31.62	37.31	-5.69	
4 & 5	3.22	16.44	-13.22	
6	-4.53	-0.35	-4.18	
7	2.25	-9.06	+11.31	*
8	-7.99	4.31	-12.3	
9	79.84	88.61	-8.77	
10	31.87	37.31	+0.56	*
5 & 6	24.56	16.44	+8.12	*
7	21.04	16.44	+4.6	*
8	26.98	16.44	+10.54	*
9	102.21	88.61	+13.6	*
10	65.16	37.31	+27.85	*
6 & 7	-9.36	-0.35	-9.01	
8	2.86	4.31	-1.51	
9	88.42	88.61	-0.19	
10	43.90	37.31	+6.59	*
7 & 8	1.36	4.31	-2.95	
9	95.78	88.61	+7.17	*
10	42.31	37.31	+5.0	*
8 & 9	92.62	88.61	+4.01	*
8 & 10	19.06	37.31	-18.25	
9 & 10	114.91	88.61	+26.3	*

* Indicates those pairs where there is a positive incremental change in NPV

Table 2.4: Positive Interactions between two projects

	Individual NPVs:		
	Both positive	One positive	Neither positive
4 + 7			✓
4 + 10		✓	
5 + 6		✓	
5 + 7		✓	
5 + 8	✓		
5 + 9	✓		
5 + 10	✓		
6 + 10		✓	
7 + 9		✓	
7 + 10		✓	
8 + 9	✓		
9 + 10	✓		

Five of the positive interactions are between projects where both have a positive NPV, that is projects 5, 8, 9 and 10. When two projects, each of which has a positive NPV are combined, we would not be surprised that their combined NPV is greater than that of the better single project. This is then the expected outcome. While projects 8 and 10 also both have positive NPVs they are strong substitutes so a positive interaction is neither found, nor expected.

Six show a project that is negative alone, namely projects 4, 6 or 7, having a positive impact on an already positive project. Two of these may be explained easily in that projects 4 and 10 affect the inner ring road, while projects 5 and 6 affect adjacent links. We might therefore expect some cumulative benefit. The remainder cannot be simply explained, and would not have been obvious a priori, illustrating the importance of a comprehensive test of pairs.

With projects 4 + 7 two negative projects interact to produce a positive result. Although this may appear odd at first, it can be seen above that projects 4 and 7 interact positively with other projects so it is not unlikely that they should interact positively with each other.

(iv) Use the information from the pairs to determine any tests of 3 project runs; if all relevant pairs show a positive interaction then the triple should be tested.

eg, 4 + 7, 4 + 10 & 7 + 10 are all positive

therefore select 4 + 7 + 10

The triple project combinations selected in this way are:-

4 + 7 + 9	5 + 8 + 9
5 + 6 + 10	5 + 9 + 10
5 + 7 + 9	7 + 9 + 10
5 + 7 + 10	

This rule yields seven three-way tests.The reason we only select these triples is that once two projects are shown to be substitutes we are no longer interested in putting them together with other projects. Higher combinations eg, quadruples should also be tested, if they can be constructed from positive pairs, in this case there were none.

In summary

(i) Test each project individually - 10 runs
(ii) Test all projects together - 1 run
(iii) Test every pair of projects - 45 runs
(iv) Test any higher combination, eg triples, that contain only positive pairs - 7 runs

Thus, a systematic approach, applied to this particular set of ten infrastructure projects requires a total of 63 runs.

At this stage the data is used to construct the first statistical model. The GLIM package (Royal Statistical Society, 1992) was used for the statistical modelling as it is interactive and quick.

Interaction terms can be specified for any combination of projects and tested for significance. The best model identified using this data is a linear regression on a set of dummy variables. It includes eight single projects, ten two-project interaction terms and one three-project interaction term, giving nineteen variables in all. Variables were omitted if the absolute size of their coefficient estimates were below four as this indicates a very small contribution to NPV. Standard errors on each were also used to guide model construction, with highly insignificant variables being discarded. The results are discussed in more detail later in this section when the two approaches are compared.

The variables and their coefficients are given in Equation (1). Further detail, including standard errors are given in Appendix 1.

$$\begin{aligned} NPV = \; & -14.03\,P1 - 8.3\,P2 - 12.42\,P3 - 7.99\,P4 + 23.60\,P5 - 5.03\,P7 \\ & +89.79\,P9 + 41.72\,P10 - 22.48\,P810 - 16.36\,P910 - 5.61\,P17 \\ & -12.19\,P27 - 12.20\,P45 - 11.22\,P59 + 13.75\,P47 \\ & +9.72\,P79 + 11.25\,P13 - 5.81\,P210 + 10.49\,P5710 \end{aligned} \qquad (1)$$

Variables

P1 to P10	dummy variables relating to each of the single projects.
P810 to P210	dummy variables for interaction between pairs, for example, P810 will be 1 if P8 and P10 are both present, zero otherwise.
P5710	dummy variable - a three way interaction, this is 1 if projects 5, 7 and 10 are all included.

The next question to be answered is how we move from this equation to a prediction of the optimum combination. The first step is to include all projects that have a positive coefficient, (ie. P5 + P9 + P10) and then add any interaction terms that include projects 5, 9 and 10 in pairs, checking that any negative interactions do not more than offset the positive effects (in which case the lowest value project of those involved would be removed).

This gives us

$$NPV = 23.6\ P5 + 89.79\ P9 + 41.72\ P10 - 16.36\ P910 - 11.22\ P59 \qquad (2)$$

So, although the interaction terms are negative they are not large enough to outweigh the gains from having all three projects. There is still a net loss from dropping any one of them. (Note that negative interaction terms are to be expected here, as adding two positive projects together is unlikely to give an NPV greater than the sum of the two projects run individually).

The next step is to consider any projects that have a negative coefficient <u>but</u> have positive interactions with any of the projects already included that outweigh the single project impact. This process adds only P7:-

$$-5.03\ P7 + 9.72\ P79 + 10.49\ P5710$$

Thus, the predicted optimum combination is P5 + P7 + P9 + P10 solving the equation gives us:

$$NPV = 23.6\ P5 + 89.79\ P9 + 41.72\ P10 - 16.36\ P910 - 11.22\ P59 \qquad (3)$$

$$-5.03\ P7 + 9.72\ P79 + 10.49\ P5710$$

Predicted NPV = <u>£142.7M</u> above the do-minimum.

Two projects 6 and 8 were omitted from Equation (1) due to insignificant coefficients. While we would not wish to add project 8 as it is a substitute for project 10, project 6 may contribute to our optimum. Thus, we have joint optima

(i) P5 + P7 + P9 + P10
(ii) P5 + P6 + P7 + P9 + P10

Both are predicted to have the same NPV. Checking both these with PLUTO runs shows P5/P6/P7/P9/P10 to be the better with an NPV of £133.5M. We have reached the optimum combination of infrastructure projects using 65 runs of the PLUTO model.

2.2 An Intuitive Approach

This approach aims to reduce the number of model runs, by defining a data set that contains examples of 'types' of combination, for example, combinations of projects on the two ring roads, on adjacent links and on unconnected links.

It is possible even with a limited data set to make some allowance for interaction by giving the regression model information on the type of link or its location which might be expected to influence its interaction. For example, if projects are undertaken on adjacent sections of road we might expect the way in which they interact to vary from that where they are not adjacent. What follows is an example of an intuitive approach, combinations have been chosen to give some variety.

A data set was constructed to test this approach which included the following:

10	tests on single projects
1	test incorporating all 10 projects
1	test of an adjacent pair on inner links (P1 + P2)
1	test of non-adjacent inner and outer link (P2 + P7)
1	test of an adjacent link and ring road (P5 + P8)
1	test of inner and outer ring road (P9 + P10)
2	tests on substitute/alternative projects affecting in one case the same link (P4 + P5) and in the other the same ring road (P8 + P10)
1	test of three adjacent links (P1 + P2 + P3)
1	test of three detached links (P1, P4 + P7)
19	model runs in all

The independent variables included a dummy variable for each project (P1 to P10). We tested a number of dummy variables chosen to reflect the specific nature of the chosen data set. These were:-

SUB -	1 if any two projects affect the same road space and may therefore be viewed as alternatives.
ADP -	1 if any two projects affect adjacent areas of road space, eg neighbouring links or a link and a ring road.
NADP -	1 if any two projects affect non adjacent areas of road space.
RRS -	1 if any pair of projects affects both ring roads
ADT -	1 if any three projects are adjacent.
NADT -	1 if any three projects are not adjacent.

We experimented with these dummy variables and identified a number of problems. In particular, any set of projects where ADT = 1 must also contain ADP = 1; the variables are not independent (similarly NADP and NADT). Moreover, none of the dummies, with the exception of SUB, contributed to the explanatory power of the model. Their standard errors were very high relative to the estimated values of the coefficients.

Our best model is given by Equation 4 (the standard errors may be found in Appendix 1). It contains seven single project variables and SUB, giving eight variables in all.

$$NPV = -6.63\ P1 - 10.31\ P2 - 18.09\ P3 + 20.19\ P5 - 12.71\ P7 + 82.53\ P9 + 34.24\ P10 - 18.18\ SUB \quad (4)$$

To find the optimum, we again add in all positive projects:-

$$NPV = 20.19\ P5 + 82.53\ P9 + 34.24\ P10$$

The interaction variable SUB is not relevant to these projects and is negative so no more projects are added. SUB has in fact led to the exclusion of P4 and P8 (the alternatives to P5 and P10) from the model as they are insignificant in the presence of SUB.

Predicted NPV = £137M above the do-minimum.

The intuitive approach has used only 19 runs so far, a systematic addition might be to add in each project in turn, retaining it if NPV increases and removing it if NPV falls. The optimum (P5/P6/P7/P9/P10) would then be reached in 26 runs, 24 runs if alternatives are excluded.

2.3 A Non-Modelling Approach

This approach starts with a run of each project individually, ten runs. The first combination includes all projects with a positive NPV

P5 + P8 + P9 + P10

as P8 and P10 are alternative treatments, the one with the higher NPV is retained giving

P5 + P9 + P10

This combination is the starting point and would be run. To reach an optimum each project would be added in turn being retained if NPV increases and rejected if NPV falls. In this way the optimum combination (P5/P6/P7/P9/P10) would be reached in 16 runs (excluding alternatives).

2.4 Model Performance

The next step was to construct a set of PLUTO runs against which the predictive accuracy of the two models in equations (1) and (4) could be tested. Table 2.5 shows the set of ten runs that we selected, chosen to exhibit variety and to include multiple combinations. We also tested the accuracy of the predicted optimum combination for each model.

Table 2.5: PLUTO Test Set

Project Combination	NPV (£m)	Characteristics
P3 + P4	-24.7	Negative NPV and negative interaction
P5 + P9	102.2	Positive NPV and positive interaction
P6 + P8	2.9	Very little interaction
P5 + P9 + P10	128.9	'Best' triple
P2 + P3 + P4	-36.5	Adjacent links
P2 + P4 + P7	-11.7	Detached links
P1 + P2 + P3 + P5	3.2	Adjacent links
P1 + P3 + P8 + P9	69.6	Detached links + 2 RRs
P5 + P6 + P7 + P9 + P10	133.5	RRs + connecting link rds
P1 + P3 + P5 + P7 + P9	90.4	Odd numbered projects

Table 2.6 contains the predictions of each model compared with the actual NPV.

Table 2.6: Road project Systematic and Intuitive Regression Models tested against actual NPV

		Systematic Model	Intuitive Model
Test Strategy	Actual NPV	Predicted NPV	Predicted NPV
P3 + P4	-24.7	-20.4 (+4.3)	-18.1 (+6.6)
P5 + P9	102.2	102.2 (0)	102.7 (+0.5)
P6 + P8	2.9	0 (-2.9)	0 (-2.9)
P5 + P9 + P10	128.9	127.6 (-1.3)	137 (+8.1)
P2 + P3 + P4	-36.5	-28.7 (+7.8)	-28.4 (+8.1)
P2 + P4 + P7	-11.7	-19.8 (-8.1)	-23.0 (-11.3)
P1 + P2 + P3 + P5	3.2	0.1 (-3.1)	-14.8 (-18)
P1 + P3 + P8 + P9	69.6	74.3 (+4.7)	57.8 (-11.8)
P5 + P6 + P7 + P9 + P10	133.5	142.7 (+9.2)	124.2 (-9.3)
P1 + P3 + P5 + P7 + P9	90.4	86.0 (-4.4)	65.3 (-25.1)
Total Absolute deviation		45.8	101.7
Average Deviation		4.6	10.2
Model Runs used		63	19
Number of variables		19	8

Figures in parentheses give the difference between the predicted NPV and the actual NPV. (The model parameters are given in Appendix 1.)

The systematic model produces prediction with a higher degree of accuracy. The intuitive model is unable to deal accurately with multi-project combinations. The systematic model is more precise, with no predictions giving an error in excess of ± 10. The interaction terms play a significant role.

2.5 Predicting the OPTIMUM

Now let us consider which model brings us closest to an optimum.

Table 2.7: Infrastructure tests - model outcomes

Approach	Model		Number of model runs	Predicted Optimum (£m)	Actual Optimum (£m)
Systematic	(i)	P5 + P7 + P9 + P10	65	142.7	131.9
	(ii)	P5 + P6 + P7 + P9 + P10		142.7	133.5
Intuitive		P5 + P9 + P10	19	137.0	128.9
No model		P5 + P6 + P7 + P9 + P10	16	-	133.5

It is possible to reach the optimum in just 16 runs using a reasonably logical approach. An extension of the intuitive method will reach it in 24 runs.

The advantage of the systematic model is that it considers the interaction between projects, and this means;

- that strong substitutes will never be included together
- that positive interactions (complements) are identified, so that even if they have a negative NPV in isolation, they can be included in the optimum.

However, if there is a limited range of infrastructure or if this is largely determined in advance due to capital constraints or political decisions (as may well be the case in local decision making) then it may be that there are very few infrastructure combinations to be tested. If there is a large range of projects, expert guidance may be necessary at an early stage in order to avoid an excessive number of model runs.

2.6 Conclusions on Infrastructure Projects

(i) The presence of a number of non-mutually exclusive infrastructure projects poses problems for modelling, because many of the interactions are unknown a priori.

(ii) It therefore appeared advisable to take a systematic approach to the construction of the data set, although this is expensive in terms of model (in this case PLUTO) runs.

(iii) Our attempt at an intuitive approach gave a poor statistical model, but with so many PLUTO runs 'in hand' over the systematic approach it was easily possible to search around the initial optimum solution using further PLUTO runs. In this way, the true optimum may be reached with fewer PLUTO runs than with the systematic approach.

(iv) It is possible with a non-modelling logical approach to reach an optimum in only 16 runs in this case.

(v) **In the presence of more information, for example on political or capital constraints, it might be possible to reduce the number of projects considered at an earlier stage.**

(vi) **The use of regression modelling for discrete projects only is not very efficient, when compared with professional judgement for identifying optimal infrastructure programmes.**

3. CORDON PRICING

PLUTO was set up to allow six possible zone boundaries for cordon charging, these are shown in Figure 1.1.

3.1 General Data Set

Our initial regression modelling was undertaken using a large data set of 45 runs. These were undertaken in an attempt to identify an optimum based on expectations of how PLUTO would respond and testing around 'good' price and cordon combinations.

The zone boundaries ranging from 1 to 6 and fees ranging from £1 to £10, give us one continuous variable (Fee) and one discrete variable (cordon location). Our data set included the independent variables Fee, which is the fee charged, and six dummy variables, zone 1 to zone 6, to indicate the zone boundary that is in existence in each case. Only one of these dummies will be active in each case, since multiple cordons were not considered. The information in the fee and zone variables was then combined:-

Fee x Zone 1 = FZ1
:
Fee x Zone 6 = FZ6

We now have six variables, each of which defines a fee to be charged at a specific zone boundary. We then squared these variables in order to try a quadratic form.

The regression model arrived at is detailed in Section 1 of Appendix 2, and listed below

$$NPV = +13.87\ FZ1 + 23.45\ FZ2 + 36.24\ FZ3 + 45.77\ FZ4 + 30.32\ FZ5 + 30\ FZ6 - 2.49\ FZ1^2 - 2.96\ FZ2^2 - 5.3\ FZ3^2 - 5.86\ FZ4^2 - 2.77\ FZ5^2 - 2.57\ FZ6^2 \quad (5)$$

This model can be used to estimate the fee for each zone that the model predicts will produce the optimum NPV. The calculations are shown in Section 1 of Appendix 2. The model predictions for an optimum are summarised in Table 3.1. Table 3.1 requires some clarification. Column A contains the fee that the regression model predicts to be the best for that zone. Column B contains the NPV that the regression model predicts for each zone if the fees are those in column A. Column C contains the NPVs obtained from PLUTO runs when using the fees in column A.

Column D shows the difference between the predicted NPV and the actual NPV at column A fee levels. Column E shows the fee levels that we have found to perform best through trial and error, and a large number of PLUTO runs. Column F shows the difference between the predicted best fee level and the actual best fee level. Column G shows the NPVs obtained using our best estimate of the optimal fees (column E). Finally, column H, shows the difference in column NPV from using the fees predicted in our model (column A) rather than the 'best' fees (column E). Column H identifies the potential 'loss' in NPV from the misspecification of the optimum by our model.

Table 3.1: Cordon Pricing - Initial Model Predictions

Zone	A Predicted Optimum fee	B Predicted NPV for A	C Actual NPV for A	D Error in NPV prediction (C-B)	E Actual 'Best' Fee	F Error in Actual fee (E/A)	G Actual NPV for E	H Under-achievement of NPV (G-C)
1	2.80	19.3	22.3	+3.0	3.00	+0.2	26.6	+4.3
2	4.00	46.3	40.9	-5.4	3.50	-0.5	51.1	+10.2
3	3.40	61.9	59.6	-2.3	3.50	+0.1	65.5	+5.9
4	3.90	89.1	86.8	-2.3	3.60	-0.3	88.3	+1.5
5	5.50	89.2	81.0	-8.2	5.00	-0.5	87.2	+6.2
6	5.90	89.4	73.9	-15.2	4.00	-1.9	80.7	+6.8

The model performs well and appears to bring us towards the best solution (a fee of £3.60 in zone 4). Any problems appear to be related to the nature of the data set. As the original PLUTO runs were undertaken with the aim of finding the best NPV, there are a large number of runs in the data with an NPV around £85M to £89M which may explain the very similar NPV predictions for zones 4, 5 and 6. Similarly a lack of data at higher fee levels (eg above £5.00) may be causing problems. A large data set will not necessarily be useful if it contains insufficient variety.

3.2 Systematic Approach

Our next task was to attempt to produce a 'good' model given a limited number of PLUTO runs. As we already had a great deal of data and a good idea of the model form it was difficult to go back to a state of innocence of the NPV output. One solution was to ask an 'expert' what he thought the best fee would be, as he was very familiar with PLUTO and its behaviour but not with the NPV figures. This gives us a 'blind' expert test. The expert recommended that the best result would occur with a cordon in zone 3 and with a fee of about £1.50. We took this advice and, in addition, ran tests with fees at half and double this value in order to try to identify a quadratic relationship.

The expert also thought that the optimum level of fee would rise as we moved the cordon outward. The expert's estimate of £1.50 enabled us to produce the following 18 test matrix

Table 3.2: Initial 'Expert' Matrix

Zone	Fee		
	1	2	3
1	0.25	0.50	1.00
2	0.50	1.00	2.00
3	0.75	1.50	3.00
4	1.00	2.00	4.00
5	1.25	2.50	5.00
6	1.50	3.00	6.00

This inital matrix is not guaranteed to enclose the optimum and in order to test for this we developed an approach of testing around the expert prediction to see if we have enclosed an optimum - if not we would need different values and we have only used three tests to discover this rather than the eighteen implicit in the matrix. The three tests centred on £1.50 in zone 3 of 0.75, 1.50 & 3.00 are the first to be run and used to locate the curve: if it is rising or falling over all three points then a further test would be run (double top value if rising, half the lowest value if falling). Then having produced a curve use the maximum point as the middle value for the next zone. This is expressed as a formal rule in Appendix 3.

This rule may be made more efficient by testing to one side of the initial value; if the result is a higher NPV continue in that direction and do not test the other side. If the result is a lower NPV then test the other side.

We started with Zone 3 at £1.50 which yields an NPV of £41.5M. The next test we tried was to double the fee to £3.00, which yields an NPV of £61.5M. Our rule now states that as the NPV is still rising we should double the fee again, to £6.00 for our third test, the test of £0.75 is now unnecessary. In this way the following set of results is obtained for zone 3.

Zone 3 Fee = £1.50 => NPV = 41.5M
3.00 => NPV = 61.5M
6.00 => NPV = 26.3M

We can now progress to the next zone taking £3.00 as our initial value (because NPV is highest at £3.00 in the previous zone). In zone 4,

Zone 4 Fee = £3.00 => NPV = £85.1M
£6.00 => NPV = £76.7M
£1.50 => NPV = £57.8M

The NPV figures for zone 4 are uniformly higher. £1.50 gives the lowest NPV in zone 4 while in zone 3 it is £6.00. This suggests that the initial premise that the optimal fee increases as the boundary zone moves out, is indeed the case. It therefore seems sensible to use this information in deciding the fees to test in the remaining zones (this is a slightly more flexible approach than that outlined in Appendix 3). For zone 5 we increase the level of the mid-fee to be tested by £1.00, thus giving a range of £2, £4 and £8 on the doubling/halving rule.

Zone 5 Fee = £4.00 NPV = £85.3
£2.00 NPV = £65.6
£8.00 NPV = £60.6

These figures suggest that we could reduce the spread of fees tested in zone 6.

Zone 6	Fee	=	£4.00 NPV = £80.7
			£3.00 NPV = £75.8
			£7.00 NPV = £68.0

The figures so far suggest that the optimum boundary in terms of NPV occurs in Zone 4 or Zone 5, in which case we would not need to test fees in zones 2 and 1. However, there may be other reasons for selecting a boundary round zone 1 or 2, so we can continue testing in order to determine the optimum fee. We expect the optimum fee to be lower in these zones given the patterns so far, so we will use the zone 3 values again, reducing the upper fee by £1.00.

Zone 2	Fee	=	£3.00 NPV = £47.8
			£1.50 NPV = £29.9
			£5.00 NPV = £40.7

£1.50 still gives a very low value so try increasing it for zone 1

Zone 1	Fee	=	£3.00 NPV = 26.6
			£2.00 NPV = 16.8
			£5.00 NPV = 19.1

We have adapted our original 18 point matrix in order to ensure that the range of fees tested for each cordon envelop the optimum. The data set gives the regression model three observations per zone.

The new model is as follows:-

$$NPV = 13.82\ FZ1 + 26.2\ FZ2 + 36.3\ FZ3 + 45.1\ FZ4 + 37.3\ FZ5 + 35.6\ FZ6 - 1.98\ FZ1^2 - 3.6\ FZ2^2 - 5.3\ FZ3^2 - 5.4\ FZ4^2 - 3.7\ FZ5^2 - 3.7\ FZ6^2 \quad (6)$$

The calculations for optimum fee and NPV are given in section 2 Appendix 2.

The results from the new model are shown in Table 3.3 (the columns are exactly equivalent to those in Table 3.1).

Table 3.3: Cordon Pricing - New Model Predictions

A Predicted Optimum fee	B Predicted NPV for A	C Actual NPV for A	D Error in NPV prediction (C-B)	E Actual best Fee	F Error in Actual fee (E/A)	G Actual NPV for E	H Under-achievement of NPV (G-C)
3.50	24.1	19.6	-4.5	3.00	-0.50	26.6	+7.0
3.60	50.2	49.9	-0.3	3.50	-0.10	51.1	+1.2
3.40	62.1	59.6	-2.5	3.50	+0.10	65.5	+5.9
4.20	97.1	88.3	-8.8	3.60	-0.60	88.3	0
5.00	94.0	87.2	-6.8	5.00	0	87.2	0
4.80	85.5	79.9	-5.6	4.00	-0.8	80.7	+0.8

This model based on only 18 runs provides an optimum fee, which in all zones except zone 1 brings us closer to an optimum NPV than the original model which was based on 45 runs. Note that because of the flatness of the model in zone 4, fees of 3.6 and 4.2 both yield NPVs of £88.3M, the highest we could find.

Thus, even in a model with just three data points per zone, the predicted 'optimum' fee and the related NPV correspond very closely to the 'best' outcome obtained by running a large number of PLUTO tests. It would, therefore, appear that three model runs, provided two of them fairly closely bound the true optimum, will be adequate to provide a reasonable estimate of the optimum fee for any one zone, although some fine tuning may then be necessary, testing points around what appears to be the 'best' outcome. The slightly higher deviance in the outer zones is largely due to the actual curve being fairly flat on top forming a plateau. Yet the predicted optimum fares are extremely close to the 'best' values found through experimentation with PLUTO runs.

3.3 Conclusions on Cordon Pricing

(i) For a continuous variable such as fee it is possible to construct a good regression model, assuming a quadratic form, with just three model (PLUTO) runs per zone, provided that these enclose the optimum and bound it fairly closely.

(ii) A large data set of model runs will not necessarily generate a useful regression model if it contains insufficient variation.

(iii) The required model (PLUTO) runs may be specified using a formal rule (Appendix 3) or a slightly more flexible adaptation of it as used here. As was noted in the previous section, further work is required on determining how much flexibility should be allowed when supposedly using formal rules, in order to get the most cost-effective result.

4. BUS FARES AND SERVICE LEVELS

4.1 Models Tested

PLUTO allows us to set bus fares in the peak and off-peak separately. They consist of a fixed charge per boarding and a distance related charge. It is possible to introduce a flat fare by setting the distance related charge to zero. Bus service levels are determined in PLUTO by a model of the behaviour of profit maximising bus operating companies responding to market conditions and, in the batch version of PLUTO we were using, it was not possible to set precise service levels exogenously. We were however able to specify that additional buses should be added to the peak and/or off-peak, fleets and this gave us some influence over service levels.

We already had to hand a very large data set of 94 PLUTO runs where fares and/or service levels were the only items to be varied. The variables were specified in the data input as:-

PS	-	number of buses per link per hour in the peak
OPS	-	number of buses per link per hour in the off-peak
PF	-	average fare paid in the peak
DPF	-	dummy, 1 if the peak fare is flat
OPF	-	average fare paid in the off-peak
DOPF	-	dummy, 1 if the off-peak fare is flat
PKF	-	average peak bus fare, converted to a factor with eight levels
OPKF	-	average off-peak fare, converted to a factor with seven levels

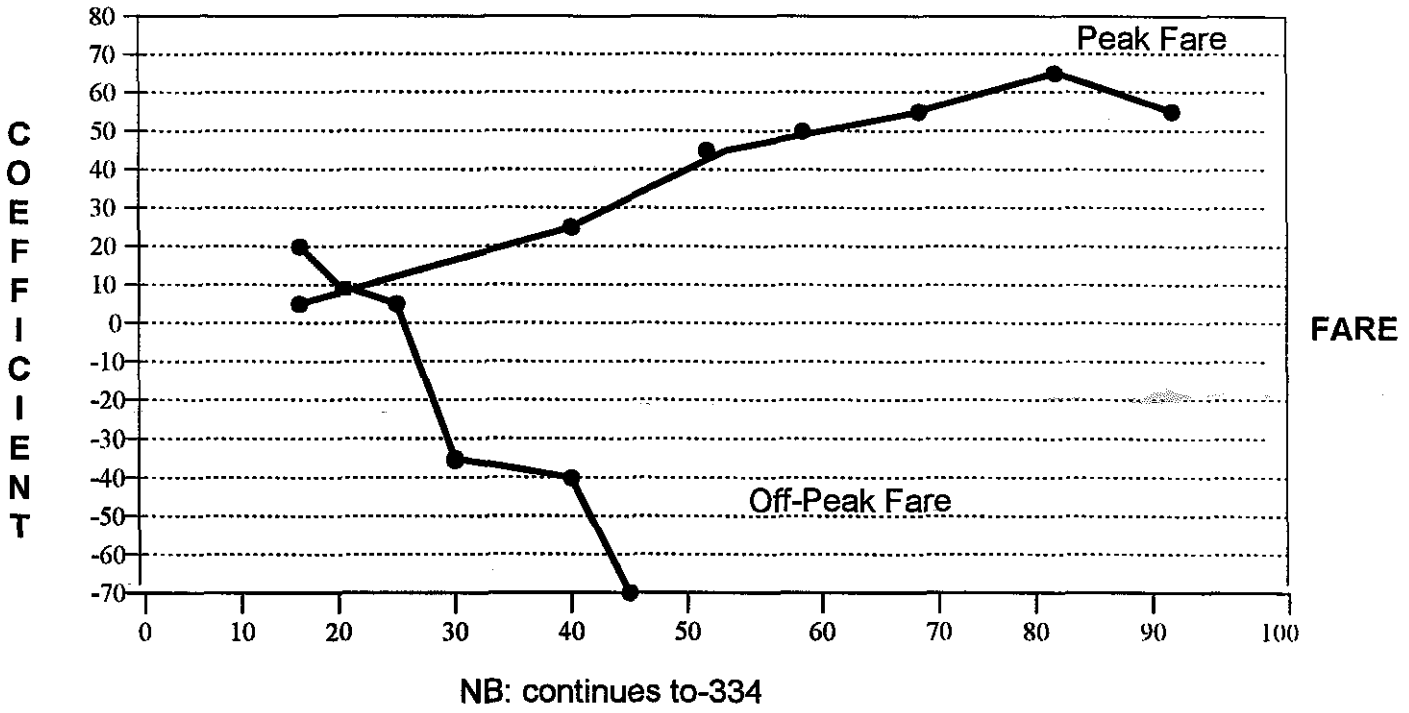
Initially we tried a quadratic by creating variables PFS (peak fare squared) and OPFS (off-peak fare squared). We also squared the service level variables giving PSS and OPSS. This model was not successful.

We next converted the fare levels to factors and used them in the model as an alternative to the continuous variables previously tried. This was far more successful. The model is shown below, further detail is given in Appendix 4.

$$\begin{aligned} NPV = & -97.75 - 4.27\,PKF2 + 27.44\,PKF3 + 44.71\,PKF4 + 50.2\,PKF5 + 55.03\,PKF6 \\ & + 62.53\,PKF7 + 55.16\,PKF8 + 17.57\,OPKF2 + 8.81\,OPKF3 + 3.6\,OPKF4 \\ & - 34.73\,OPKF5 - 41.41\,OPKF6 - 334.9\,OPKF7 - 0.08\,OPSS + 8.5\,OPS \end{aligned} \qquad (7)$$

The only variables to contribute to the model in addition to the fare factors are the off-peak service level and the off-peak service level squared. It is illuminating to graph the relationship between fare level and NPV, see Figure 4.1. The graph suggests that the peak fare may well exhibit a roughly quadratic relationship with NPV, but the vast majority of our observations are to the left of the optimum. The off-peak fare plot, clearly suggests trying some lower fares between 10 pence and zero.

Figure 4.1: Contribution of Different Fare Levels to NPV



Taking the model at face value the best peak fare found is PKF7 (78p) and the best off-peak fare found is OPKF2 (10p) although as lower fares have not been tested, it is possible that a fare of less than 10p may yield a higher NPV. The optimal off-peak service level can be calculated as approximately 53 buses per link per hour. Nothing is said about the peak service level. The predicted NPV is

$$NPV = -97.75 + 62.53 + 17.57 - 0.08(53)^2 + 8.5(53)$$

$$= 208.13$$

i.e. an increase of £208M over the do-minimum. This is largely accounted for by the improvement in off-peak service levels.

The model is telling us that although we can obtain a reasonable result from a large data set we happen to have to hand, a better result is likely to be obtained through the use of a data set where the model (PLUTO) runs have been designed with a view to providing GLIM with suitable information. In this case a large number of the runs are to the left of the peak fare optimum, while we have not enclosed the off-peak optimum.

4.2 Conclusions on Bus Fares and Service Levels

(i) The regression model suggests a high peak/off-peak fare differential.

(ii) There is a case for experimenting with off-peak fares of less than 10 pence.

(iii) The model also suggests that increasing the off-peak service level from 11 up to an average of 53 buses per hour per link will have a very large positive impact on NPV.

(iv) **Regarding flat fares we found mixed evidence and they did not appear in our preferred model.**

(v) **Working in average fares has not proved sensible. Instead of peak and off-peak fares, each with a dummy to indicate if they are flat fares, we now feel that it would be better to work with the parameters that PLUTO actually sets, namely the fixed and variable parts of the fare, each for peak and off-peak. Similarly we have had difficulty in controlling PLUTO bus service levels, since these are not determined explicitly by parameters, but can merely be influenced, with other influences as patronage and road speed changes playing their part.**

5. CORDON PRICING, BUS FARES AND SERVICE LEVELS

5.1 Initial Tests

The next step was again to return to the original problem of identifying an optimum with a minimum number of model runs. The tests reported here widened the scope of the problem by allowing policy in three areas to vary at once, namely, cordon pricing, bus fares and bus service levels. This gives us the option of change in any one of the following variables:

Cordon Pricing	-	FEE	-	Fee charged to cross the cordon in the peak
		Z1 to Z6	-	Zone boundary for cordon charge
Bus Fare	-	PFF	-	Peak fare, fixed element
		PFV	-	Peak fare, variable element
		OFF	-	Off-peak fare, fixed element
		OFV	-	Off-peak fare, variable element
Bus Service	-	APB	-	Additional vehicles in the peak
		AOB	-	Additional vehicle in the off-peak

The exact specification of the input variables will be, in part, dependent on the strategic transport model used. In this case, using PLUTO, we have some constraints on the ways in which bus service levels may be specified by the policy maker. PLUTO automatically adjusts bus service levels, as other policy variables are altered, in order to maintain a certain load factor. Moreover, policy changes to service levels are expressed in terms of additional vehicles in the peak and/or off-peak. Thus, it is very difficult to impose a specific change in the level of service provided. After some experimentation a decision was taken to allow PLUTO to set the peak service level in response to changes in demand; the option of adding vehicles in the off-peak was retained.

For the first experiment an 18 test matrix was devised, designed to provide the minimum number of runs that would permit the construction of the first regression model. The intention was to use the results of this initial regression model in order to suggest further PLUTO runs that would both add to the regression data set and move us towards an optimum. The original 18 test matrix is shown in Table 5.1.

Table 5.1: Formal 18 Test Matrix for Bus Service and Fare Levels and Cordon Pricing tests

Run Number	Run Code	Cordon Pricing		Peak Bus Fares		Off-peak Bus fares		Additional off-peak vehicles
		Zone	Fee	Fixed	Variable	Fixed	Variable	
1	LM	Z1	0.4 CE1	2PFE	2PVE	PFE	PVE	0.6BP
2	LO	Z2	0.4 CE2	PFE	0.6PVE	0.6PFE	0.3PVE	2BP
3	LP	Z3	0.4 CE3	3PFE	PVE	0.6PFE	0.3PVE	BP
4	LQ	Z4	0.4 CE4	0.6PFE	0.6PVE	0.6PFE	0.6PVE	BP
5	LS	Z5	0.4 CE5	PFE	PVE	0	0.3PVE	0.6BP
6	LT	Z6	0.4 CE6	2PFE	0.6PVE	2PFE	0.6PVE	2BP
7	LU	Z1	CE1	PFE	PVE	PFE	PVE	0.3BP
8	LV	Z2	CE2	3PFE	3PVE	PFE	PVE	3BP
9	LW	Z3	CE3	PFE	0.3PVE	0.6PFE	0	BP
10	LX	Z4	CE4	2PFE	0.6PVE	2PFE	0	BP
11	LY	Z5	CE5	2PFE	0.3PVE	2PFE	0	0.3BP
12	LZ	Z6	CE6	0	5PVE	0	5PVE	3BP
13	KQ	Z1	2CE1	2PFE	PVE	0.6PFE	0	0
14	KR	Z2	2CE2	0	PVE	0	0.6PVE	5BP
15	KS	Z3	2CE3	9PFE	3PVE	0	2PVE	BP
16	KT	Z4	2CE4	5PFE	5PVE	5PFE	5PVE	BP
17	KU	Z5	2CE5	0.6PFE	0.3PVE	0.6PFE	0.3PVE	0
18	KV	Z6	2CE6	0	0.3PVE	0	0	5BP

Notes

(i) CE1 to CE6 are based on 'expert' advice on fees for the six zones. When setting these values it is possible to take account of advice as to whether fees should rise, fall or remain constant as the boundary moves out from the centre. If no expert advice is available, guidance might be sought from previous studies, or a random pattern used.

(ii) PFE and PVE are values derived from expert advice or do-minimum values for the peak fixed and variable fare elements respectively. They are used for both peak and off-peak fares.

(iii) BP is a proportion of the number of buses currently used in the peak. When using PLUTO we adjust the off-peak only.

The next task was to specify the starting values for each variable. We used a combination of common sense constraints and limited expert guidance.

Cordon Price and Zone (CE1 to CE6) - Earlier in the work advice from an 'expert' was that a 'good' fee would be around £1.50, and that it was likely to increase as the boundary moved out. Our starting value for the cordon fee was therefore set at £1.50 to allow us to test around it. At this stage we had no firm guidance as to the best location for a zone boundary so we tested six possibilities, forming concentric rings around the central area. This appeared to be a reasonable starting point as no UK city has a cordon charge as yet, knowledge is bound to be limited. As the boundary moves out, the fee increases in steps of 50 pence, in order to ensure that a wide range of values is tested.

Bus Fares (PFE, PVE) - Bus fares in the PLUTO do-minimum run are frozen at the commercial levels prevailing in year 1; that is PFF = 40p, PVF = 10p, OFF = 20p and OFV = 5p. By year 10 in the do-minimum a small amount of subsidy is being paid. A common sense constraint was to keep fares at or below the peak commercial fare level, as justification of a higher fare level would be problematic. However, we do have some tests of higher fare levels to ensure sufficient variation and to avoid missing the possibility of a successful high fare strategy. It was also seen to be sensible to constrain the off-peak fare to be equal to or lower than the peak fare, to reflect the different price elasticities. With these constraints in mind our starting values are PFE = 20p and PVE = 5p, which also happen to be the level of off-peak fares in the do-minimum run.

Bus Service Levels (BP) - The constraint here, again a common sense one, was that we did not wish to test service levels below commercial levels. It was intended to add additional off-peak vehicles. Our original unit here is 10% of the peak vehicle requirement (312) rounded to the nearest 10 to give BP = 30.

Table 5.2 shows the test matrix values and the NPV resulting when each strategy was run using PLUTO.

Table 5.2: Cordon Pricing, Bus Fares and Service levels the 18 test matrix values (all prices in pence)

Run Number	Run Code	Cordon Pricing		Peak Bus Fare		Off Peak Bus Fare		Additional Off pk buses	NPV (£m)
		Zone	Fee	Fixed	Variable	Fixed	Variable		
1	LM	1	60	40	10	20	5	18	12.02
2	LO	2	80	20	3	12	2	60	125.12
3	LP	3	100	60	10	12	2	30	120.25
4	LQ	4	120	12	3	12	3	30	102.14
5	LS	5	140	20	5	0	2	18	97.55
6	LT	6	160	40	3	40	3	60	64.81
7	LU	1	150	20	5	20	5	9	-2.85
8	LV	2	200	60	15	20	5	90	125.01
9	LW	3	250	20	2 *	12	0	30	136.55
10	LX	4	300	40	3	40	0	30	142.67
11	LY	5	350	40	2	40	0	9	102.86
12	LZ	6	400	0	25	0	25	90	-172.98
13	KQ	1	300	40	5	12	0	0	85.02
14	KR	2	400	0	5	0	3	150	185.14
15	KS	3	500	180	15	0	10	30	-62.99
16	KT	4	600	100	25	100	25	30	-571.03
17	KU	5	700	12	2	12	2	0	46.89
18	KV	6	800	0	2	0	0	150	18.54

* All values of 2 pence have been rounded up from 1.5 pence.

This data was then used to construct a linear regression model. Test 16 has a very large, negative NPV, a result of the combination of very high bus fares with a high cordon charge. We found it impossible to produce a viable regression model with this run included. We omitted it on two grounds, firstly, that the NPV was beyond our range of interest, secondly, the distorting effect it had on the statistical model.

The data set was limited to 17 observations, which constrained the number and combinations of variables that we were able to test within the regression model. However, at this stage we are looking to the regression model for guidance on additional PLUTO runs rather than a definitive answer. The best outcome was obtained using the model listed below in Table 5.3. It should be remembered that only one cordon charge boundary will be operating at any one time; so only one of FZ1 to FZ6 will be relevant.

Table 5.3: Initial GLIM Model (8) using combined data (17 observations)

	Estimate	S.E.	Parameter	Variable Name
1	-510.3	731.0	1	Constant
2	40.22	43.48	FZ1	
3	-19.66	70.61	FZ2	Fee x Zone*
4	35.23	62.67	FZ3	
5	52.32	74.74	FZ4	
6	45.97	67.71	FZ5	
7	29.05	61.37	FZ6	
8	1.906	1.976	APF	Average Peak Fare
9	2.084	1.363	AOF	Average Off-peak Fare
10	-7.290	2.582	FEES	Fee squared
11	8.886	29.63	PS	Peak service level
12	18.69	13.73	OS	Off-peak service level
13	-0.06409	0.01423	AOFS	AOF - squared
14	-0.006596	0.04204	APFS	APF - squared

$R^2 = 99.5$

* Z1 to Z6 are a series of dummy variables, 1 for each zone, taking the value if the boundary is in that zone, zero otherwise

We used this model to aid in specifying six further PLUTO runs. Table 5.4 shows the levels of the input variables and the NPV obtained from the runs. The fees to be charged in each zone were derived directly from the model by differentiation (except in the case of zone 2 when the coefficient is negative). Guidance on the other variables is more difficult to extract, but a higher level of off-peak service is indicated. A high peak fare, above commercial levels, is also suggested, as is a lower off-peak fare.

Table 5.4: An additional six PLUTO runs (values in pence)

	Inputs							Output
Code	Zone	Fee	Peak Fare Fixed	Peak Fare Variable	Off Peak Fare Fixed	Off Peak Fare Variable	Additional O/P Vehicles	NPV (£M)
KZ	1	300	80	20	0	5	200	154.96
JA	2	300	40	10	10	0	120	238.06
JB	3	250	40	0	15	0	100	222.13
JC	4	370	20	10	0	2	250	277.39
JD	5	330	60	15	10	5	60	189.53
JE	6	200	20	20	20	0	120	240.19

As can be seen from the NPVs, we are heading in the right direction. This data was added to the data set and the regression model rerun. Table 5.5 shows the best performing regression model obtained.

Table 5.5: Combined data regression model (9) 23 observations

Estimate	S.E.	Parameter	Variable Name
-37.21	16.16	1	Constant
1.517	0.2410	APF	Average Peak Fare
1.033	0.5311	AOF	Ave off-peak Fare
-0.008649	0.001172	APFS	Ave Peak Fare Squared
-0.06408	0.005851	AOFS	Ave Off-peak Fare Squared
38.91	6.546	FZ1	
47.27	7.172	FZ2	Fee x Zone for each zone
57.29	6.855	FZ3	
72.01	6.622	FZ4	
58.94	6.818	FZ5	
66.46	8.874	FZ6	
-7.064	0.9602	FEES	Fee squared
1.605	0.2158	AOP	Additional Off-peak vehicles
-0.005148	0.0009432	AOPS	AOP squared

Deviance 1084.2 (original deviance 234218) $R^2 = 99.5$

Although we were able to experiment with a larger range of variables that we thought would contribute to the model, eg changing FEES into six zone specific variables, adding a term linking the relative levels of peak and off-peak fares, these were not successful. The interaction variables were insignificant, as were four of the six fee squared variables. Moreover, the six fee squared variables, by reducing the degrees of freedom, made the model as a whole less reliable. The only new variable tried that performed well was the squared form of the additional off-peak vehicles variable.

This model, far from perfect though it is, was then used to predict an optimum PLUTO run (JG). The calculations are shown in Appendix 5. The PLUTO run suggested by regression model (9) is shown together with the "best" PLUTO run from the input data set, in Table 5.6.

Table 5.6: 23 data points, combined data, predicted optimum.

	Predicted Optimum (JG)	Best Output (JC)
Zone boundary	4	4
Fee	£5.10	£3.70
Peak Fare - Fixed	0.40	0.20
- Variable	0.12	0.10
Off-peak - Fixed	0.08	0.00
Variable	0.00	0.02
Additional off-peak vehicles	156	250
NPV (PLUTO Output)	279.2M	277.39M

The optimum run suggested by model 9 gives an actual PLUTO outcome which exceeds the values obtained in all the input data tests.

The two strategies are very similar in many respects, with our "optimum" run setting a higher road user fee and peak bus fare. In JG annual income from road user fees exceeds, by £3M in year 10, the subsidy to the bus company, while in JC the revenue from road user fees is exceeded by the subsidy to the bus company by about £5m in year 10. The policy in JC is running a revenue account deficit, while JG manages a slight surplus.

5.2 Cordon Pricing, Bus Fares and Bus Service Levels - Interim CONCLUSIONS

In this section we have tackled an extremely difficult problem and have had considerable success in reaching an optimum. 24 model runs were undertaken, with the starting levels for each input variable based on a mixture of common sense constraints and limited expert advice.

(i) We have now shown that our method is able to predict the input levels that will take us towards an optimum, while allowing a large number of inputs to vary at once.

(ii) The problem set in this section is a very difficult one given the number of possible zone boundaries and the different elements of the bus fares.

(iii) If the range of options could be narrowed a little, through further use of expert advice (boundary setting), simplification (fares structure) or constraints as to what is or is not feasible, the problem would become even more tractable.

5.3 Revised Tests

The tests above were carried out rapidly. When we came to consider how one would make concrete recommendations as to the method, it was clear that the original model (8) shown in Table 5.3 would be difficult to identify following a set procedure as it includes a number of insignificant variables. In addition, we had identified a new variable, AOPS (additional vehicles in the off-peak, squared) which was very useful in the second model. It was decided to go back and test it in the original model, as this would enable us to specify the "optimum" test at this stage also, as every input variable would have a quadratic form in the regression model. It was also felt that the regression modelling aspect would be more easily specified as a "procedure" to be followed if the initial model was tidied up a bit. The new model for the original 17 item data set is as follows;

Table 5.7: Revised regression model (10) for combined data, 17 observations

Estimate	Standard Error	Parameter	Variable Name
-76.94	45.11		Constant
1.482	0.3643	APF	Average peak fare
1.182	0.7027	AOF	Average off-peak fare
-0.008044	0.001714	APFS	APF squared
-0.06679	0.008033	AOFS	AOF squared
56.04	16.61	FZ1	Fee * Zone for each zone
1.892	59.11	FZ2	
57.17	10.09	FZ3	
73.94	9.914	FZ4	
70.84	13.22	FZ5	
50.26	30.99	FZ6	
-7.989	1.493	FEES	Fee squared
2.804	1.185	AOP	Additional off-peak vehicles
-0.002961	0.006363	AOPS	AOP squared

Deviance 579.13 (original deviance 132266) $R^2 = 99.6$

This was the best regression model we could find using this data set. While it is not ideal, it is better than the original model shown in table 5.3. The are several points to note, firstly, FZ2, is positive although low and with a high standard error, in the original it had a negative coefficient and an even higher standard error. Secondly, the constant has been reduced significantly. Thirdly, and most importantly, we now have a model that will allow us to make a first estimate of the "optimum". If we use differentiation, as before, we come up with suggested values as follows;

92p Average peak fare
9p Average off-peak fare
4 Zone
470p Fee
473 Additional off-peak vehicles

The AOPS variable is insignificant so there is some doubt over the very large value for additional off-peak vehicles. The data set thus far probably does not contain values of AOP high enough for the contribution to NPV to start declining or even slowing significantly. It would be our recommendation that the squared variables are retained in order to allow initial estimates of the optimum values to be made. These values were then substituted into model 10

$$\begin{aligned} NPV &= -76.94 + 0.3643(92) + 1.182(9) - 0.008044(92)^2 - 0.06679(9)^2 + 73.94(4.7) - 7.989(4.7)^2 \\ &\quad +2.804(473) - 0.002961(473)^2 \\ &= -76.94 + 33.5156 + 10.638 - 68.084 -5.41 + 347.518 - 176.477 + 1326.292 - 662.46 \\ &= £728.59M \end{aligned}$$

The values were then input into a PLUTO run (JH), which gave an NPV of £231.12M, which is significantly higher than the best (KR) obtained in the input data set, see Table 5.8. The regression model prediction of £728.59M is quite inaccurate at this stage, however, the model has succeeded in guiding us towards a strategy with a higher NPV.

Table 5.8: Predicted optimum from 17 data points revised model (values in pence)

Input variable	"optimum" run (JH)	Best of input data set (KR)
Peak fare - fixed	40	0
- variable	13	5
Off-peak fare - fixed	9	0
- variable	0	3
Additional off-peak vehicles	473	150
Zone boundary	4	2
Fee	470	400
NPV	£231.12M	£185.14M

Our optimum run succeeds by offering a very good service to off-peak bus users and by speeding up peak traffic flows. We have obtained a good result on the 17 point data set. The question now is how far we can improve this with additional PLUTO runs.

The GLIM model was used to guide us in setting up further PLUTO runs. Firstly we estimated the "best" fee for each zone, again using differentiation.

Zone 1	£3.50
Zone 2	£0.10
Zone 3	£3.60
Zone 4	£4.70
Zone 5	£4.50
Zone 6	£3.20

These fees were then included in a matrix of six additional tests, except where the values are close to previous runs already in the data set. We also increased the fee on Zone 2 slightly, as this variable is insignificant in the model. Values for other input variables were designed to test around the values predicted to perform best by the regression model. Previous PLUTO runs in the existing data set were also used to guide the design, in order to avoid replication and ensure variation in the data set. The resulting six tests are shown in Table 5.9.

Table 5.9: An additional six test matrix (all prices in pence)

Code	Zone	Fee	Peak Fare Fixed	Peak Fare Variable	Off peak Fare Fixed	Off peak Fare Variable	Additional off-peak vehicles	NPV £M
JJ	1	400	60	5	8	0	400	186.61
JK	2	40	45	10	5	0	600	91.0
JL	3	360	25	10	0	2	200	245.06
JI	4	400	60	5	2	1	350	274.07
JM	5	500	10	20	0	0	150	266.88
JN	6	240	20	15	2	0	500	187.38

The six PLUTO runs above and the "optimal" run were added to the original data set, giving us 24 observations in total. The regression model was then reestimated.

Table 5.10: New data set regression model (11), 24 observations

Estimate	Standard error	Parameter	Variable Name
11.02	34.04		Constant
22.96	13.04	FZ1	
38.27	15.03	FZ2	Fee * Zone for each zone
39.39	14.08	FZ3	
47.37	13.16	FZ4	
44.18	13.98	FZ5	
51.52	17.34	FZ6	
-5.532	1.843	FEES	
1.259	0.5188	APF	Fee squared
0.8808	1.047	AOF	Average peak fare
-0.007174	0.002349	APFS	Average off-peak fare
-0.0579	0.01233	AOFS	APF squared
0.84	0.1971	AOP	AOF squared
-0.00145	0.000321	AOPS	Additional peak vehicles AOP squared

This model is again fairly successful ($R^2 = 0.976$); all the variables have the expected signs, and most are significant. This model points us towards zone 6 rather than zone 4, but as they are close it might be wise to test both as the standard errors are high. Again we need to estimate the optimum input variables that the model is suggesting;

Fee = 51.52/11.064 = 4.656 = £4.70
APF = 1.259/0.014348 = 87.7 = 88p
AOF = 0.8808/0.1158 = 7.6 = 8p
AOP = 0.84/0.0029 = 289.6 = 290

The regression fee for zone 4 was also estimated since its coefficient was clearly not significantly worse than that for zone 6, and zone 4 was the predicted optimum zone boundary in model 10.

Fee = 47.37/11.064 = 4.28 = £4.30

The regression model prediction for NPV using these inputs with the cordon around zone 6 was calculated;

$$\begin{aligned} NPV &= 11.02 + 51.52(4.7) - 5.532(4.7)^2 + 1.259(88) + 0.8808(8) - 0.007174(88)^2 - 0.0579(8)^2 \\ &\quad + 0.84(290) - 0.00145(290)^2 \\ &= 11.02 + 242.144 - 122.20188 + 110.792 + 7.0464 - 55.55 - 3.7056 + 243.6 - 121.945 \\ &= £311.2M \end{aligned}$$

The optimum combination of inputs predicted was then tested using PLUTO; for zone 6 (JO) and for zone 4 where the regression model predicted an NPV of £292.66M

Table 5.11: Optimum and the best of the data set.

Variable	Optimum (JO)	Optimum - zone 4	Best of the data set (JM)
Peak fare - fixed	39	39	60
- variable	13	13	5
Off-peak fare - fixed	8	8	2
- variable	0	0	1
Additional off-peak vehicles	290	290	350
Zone boundary	6	4	4
Fee	470	430	400
NPV (PLUTO output)	£274.67M	£288.35M	£274.07M

Again we have obtained an NPV higher than any in the input data set, using zone 6 as suggested by the regression model, however, by also testing zone 4 which had a coefficient only slightly less than that of zone 6, we obtain an even higher value for NPV. The regression model prediction is considerably more accurate in this model at £311.2M although still 8% higher than the true value.

An additional six model runs have obtained an increase in NPV of approximately £50M.

A slight problem was experienced with the additional six tests as the outcome fare levels in each case were below or equal to the predicted best value. However, it is not thought that this has influenced the result greatly.

5.4 Conclusions on Bus Fare, Service Level and Cordon Pricing Tests

(i) It is very useful if a quadratic form can be approximated for each continuous variable, as this allows predictions of optimum levels for such variables even where the statistical model is far from perfect.

(ii) The statistical models constructed may be imperfect, but so long as they are the best attainable with a given data set, they should point in the direction of the answer.

(iii) Using only 24 model runs we have attained a high NPV, faced with a large number of variables.

(iv) Moving from 17 to 24 model runs obtained an increase in outturn NPV of 24.7%.

6. INTRODUCING FURTHER COMPLICATIONS

6.1 AN ADDITIONAL INDEPENDENT VARIABLE

So far we have attempted to optimise while allowing the policy to vary on the following: bus fares, cordon pricing and bus service levels. At this point we add the option of a predefined package of infrastructure investment to the other three policy options. We have already undertaken a considerable amount of analysis to discover the best combination of road infrastructure projects, see Section 2, and in the previous section we were moving towards an optimum based on fares, service levels and cordon pricing. In designing a six test matrix for PLUTO, we took our previously defined best combination of infrastructure projects, as our infrastructure package, and varied the level of the other variables around what the regression models tell us is their optimum level. The resulting matrix is given in Table 6.1 below;

Table 6.1: 6 Additional PLUTO tests including infrastructure

Code	Zone	Fee	Peak Fare Fixed	Peak Fare Variable	Off peak Fare Fixed	Off-peak Fare Variable	Extra off peak vehicles	Infra-structure package included	NPV £m
JQ	1	500	50	0	10	0	700	YES	222.15
JR	2	600	30	10	6	1	200	YES	383.21
JS	3	300	0	20	0	2	300	YES	380.68
JT	4	450	100	0	20	0	400	YES	402.19
JU	5	550	15	20	3	2	600	YES	281.12
JV	6	700	0	40	0	4	500	YES	270.52

This data was then added to the 24 point data set used in Section 5.3, to give us thirty observations. The infrastructure package was a dummy variable, 0 = no investment, 1 = projects 5,6,7,9,10. The regression model was rerun using the same variables as before, as shown in Table 6.2.

Table 6.2: Regression model (12), 30 observation data set

Estimate	Standard Error	Parameter	Variable Name
30.0	37.94		Constant
31.99	14.17	FZ1	Fee x Zone for
45.03	16.48	FZ2	each zone
45.31	16.00	FZ3	
50.40	15.03	FZ4	
41.17	15.43	FZ5	
45.55	18.52	FZ6	
-4.896	2.080	Fees	Fee squared
1.081	0.5686	APF	Average peak fare
0.05693	1.203	AOF	Average off-peak fare
-0.0073	0.0027	APFS	APF squared
-0.04435	0.0133	AOFS	AOF squared
0.5589	0.1861	AOP	Additional off-peak vehicles
-0.00092	0.00027	AOPS	AOP squared
144.3	2.08	INF	Infrastructure package

This is a useful model, giving the expected signs on all the coefficients, and showing the beneficial impact of the infrastructure package. The coefficient for the infrastructure package is actually higher, though very similar to the NPV obtained from running this investment package through PLUTO in isolation, £133.5m (see section 2).

Differentiation was used to obtain the values for each variable that the model implied would enable us to reach an optimum;

74p Average peak fare
0.6p Average off-peak fare (entered as a zero fare)
515p Fee in zone 4
305 Additional off-peak vehicles
1 Infrastructure package

Comparing these figures with those for the previous model without the infrastructure package, it suggests slightly lower bus fares, a few more off-peak vehicles, and a higher cordon charge, this time in zone 4. The marginal nature of most of the changes suggests that the regression model is now fairly stable. When these values were input into PLUTO the NPV obtained was £295.63M, which is higher than all but one in the data set, suggesting that even with only six runs including the infrastructure package, a helpful result can be obtained.

6.2 BEHAVIOUR OF OTHER DEPENDENT VARIABLES

In this study we have concentrated on optimisation with NPV as the variable to be maximised, as NPV is the most complete measure of the costs and benefits of a particular policy combination. Of course some impacts are not - at present - monetised, most particularly environmental impacts which are not included in our NPV at all. It might also be an objective to reduce accidents, in which case the absolute number becomes a decision variable, regardless of the fact that the cost of accidents is incorporated in the NPV measure.

In this section we consider how environmental impacts and injury accidents vary as we move toward an optimum NPV. Here we will use the data set built up in Sections 5 and 6.1 to investigate the issue. The environmental index represents various types of air and noise pollution; the lower the number the less the pollution. Accidents are shown as total casualty accidents per year. Table 6.3 shows the three dependent variables of interest, NPV, accidents and the environmental index, together with the policy variables for each run. The data has been ordered in relation to NPV, listing the runs with the highest NPV first. It is clear from the table that maximising NPV has also led to environmental improvements and a reduction in injury accidents relative to less successful runs in terms of NPV.

The highest NPVs are generally found with low off-peak bus fares, higher off-peak service levels, a high cordon charge and with investment in road improvements; the net result is fewer car journeys to the centre and hence fewer accidents and reduced pollution. The environmental indicator is well correlated with NPV at least in the context of this particular set of policy variables. The accident figure is less well correlated; the lowest level is achieved by a policy of very low bus fares 8p in the peak and zero in the off-peak and the highest congestion charge - £8 - tested, set at the boundary furthest from the centre; thus reducing traffic flow into the centre. Such a policy cannot be expected to maximise NPV as it involves running a large subsidy to the bus company, and time or money penalties to car drivers (switchers and stayers). A policy involving a severe reduction in traffic could only succeed overall if coordinated with policy options, such as land use changes, cycle lanes, pedestrianisation, which we have not examined here.

Table 6.3: NPV, Accidents, Environmental Index and Policy Variables

NPV (£m)	Environmental Index	Accidents	Ave Pk Fare (p)	Ave O/Pk Fare (p)	Additional Vehicles	Zone Boundary	Fee (p)	Infrastructure
402.19	47.6	1619	100	20	400	4	450	1
383.21	**47.2**	1604	68	10	200	2	600	1
380.68	47.6	1581	73	8	300	3	300	1
281.12	47.8	1570	90	11	600	5	550	1
274.07	48.8	1587	79	6	350	4	400	0
270.52	48.1	1632	143	16	500	6	700	1
266.88	48.7	1571	84	0	150	5	500	0
245.06	49.0	1617	84	8	200	3	340	0
231.12	49.0	1584	89	9	473	4	470	0
222.15	49.4	1724	50	10	700	1	500	1
187.38	50.2	1699	75	2	500	6	240	0
186.61	50.1	1795	79	8	400	1	400	0
185.41	49.2	1641	19	12	150	2	400	0
180.54	48.2	**1521**	8	0	150	6	800	0
142.67	51.0	1722	42	40	30	4	300	0
136.55	50.0	1667	28	12	30	3	250	0
125.12	50.5	1712	31	20	60	2	80	0
125.01	50.9	1782	116	40	90	2	200	0
120.25	51.2	1781	97	20	30	3	100	0
102.86	50.7	1697	48	40	9	5	350	0
102.14	50.8	1730	23	24	30	4	120	0
97.55	50.4	1719	39	8	18	5	140	0
91.00	51.1	1734	83	5	600	2	40	0
85.02	51.2	1885	60	12	0	1	300	0
64.81	51.9	1822	52	52	60	6	160	0
46.89	49.0	1598	20	20	0	5	700	0
12.02	52.6	1938	78	40	18	1	60	0
2.85	52.3	1905	39	40	9	1	150	0
-62.99	50.7	1788	236	39	30	3	500	0
-172.98	52.3	1827	90	93	90	6	400	0

7. CONCLUSIONS

7.1 Discrete Policy Variables

Our work with discrete variables, ie a number of non-mutually exclusive infrastructure projects (see Section 2) poses problems for modelling, where many of the interactions are unknown a priori. A fully systematic approach to such a problem is expensive in terms of transport model runs. While an intuitive approach may be unable to reach the optimum, it requires fewer transport model runs, allowing for some testing around the predicted optimum. However, an approach based more simply on professional judgement is probably more efficient.

7.2 Continuous Policy Variables

When investigating a single continuous policy variable, tests as described in Section 3 to identify the area of the optimum NPV through the assumption of a roughly quadratic form proved to be very effective.

7.3 Multiple Policy Variables

When considering a wide range of policy variables the design of the data set should aim to include as much variation as possible for each input variable. Although a pure orthogonal design would require too many transport model runs, the same principles should guide the design process.

The main points are:-

(i) For the initial data set, select the minimum number of runs that gives variation in the variables under consideration, and gives sufficient degrees of freedom to test regression models with the policy variables (and squared terms for continuous variables) as independent variables,

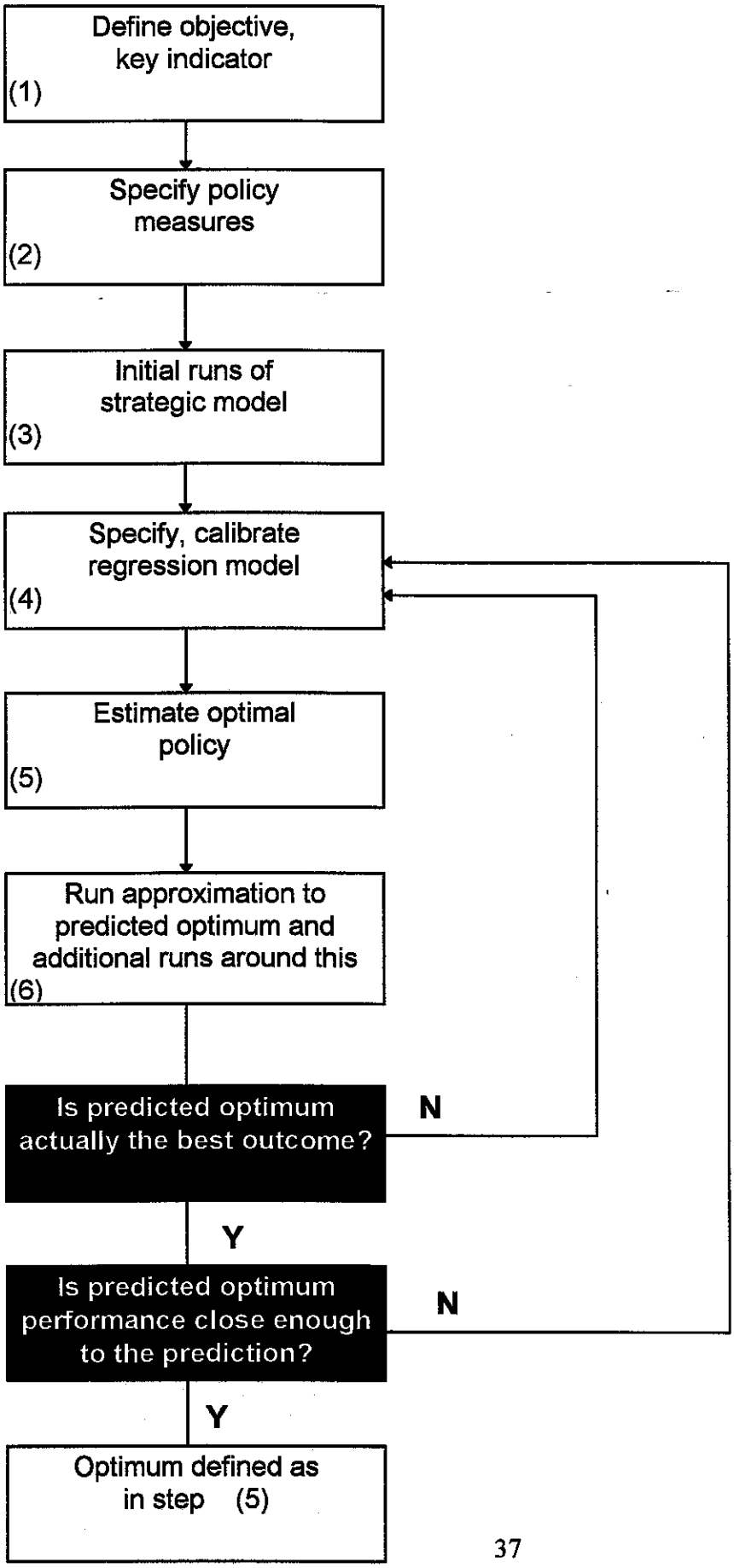
(ii) use regression models to guide the selection of further transport model runs.

This process should continue until improvements in NPV become marginal and exactly what 'marginal' means in the context in which the process is being used should be defined in advance.

7.4 The Optimisation Process

Drawing on our experimental work with PLUTO we have developed a general methodology for the optimisation process which may be expressed as a six step process with iteration, this is summarised in diagrammatic form in Figure 7.1.

Figure 7.1 : The optimisation process



The steps to be followed:

1. **Define the objective function and select a key indicator for that objective.**
   In the work reported here our objective is economic efficiency and the key indicator the NPV. We chose to use NPV as our indicator in the developing and testing process as it is an important indicator in many transport appraisals and because clear optima can be identified.

2. **Specify the policy measures to be tested.**
   This will involve the identification of policies with the potential to contribute to a successful strategy. It is also important to define a variable or parameter for each policy to enable it to be included in the regression model.

3. **Initial runs of the strategic model.**
   Sufficient runs are required to generate an initial regression model. The number of runs needed will be dependent on; the number of independent variables and any interaction terms between them and; by the need to provide a range of policy options.

4. **Specify and calibrate a regression model.**
   The key indicator (step 1) forms the dependent variable, and the policy variables (step 2) the independent variables. This is an exploratory stage which should yield information on the usefulness of policies and the presence or absence of synergy through interaction terms.

5. **Estimate optimal policy.**
   Interpret the regression model to estimate the optimal value of the key indicator and the set of policy variables required in order to achieve it.

6. **Run approximation to predicted optimum and additional runs around this.**
   The strategic model is used to run a policy test closely approximating the predicted optimum found in step five, together with other runs around this optimum. The number of such runs, and their closeness to the predicted optimum will depend on the goodness of fit of the regression model, particularly as expressed through t statistics and residual values. Where the model fits very well, only the predicted optimum plus a couple of close by tests need be run. Where the model remains poorly fitting, rather more runs are indicated, and spread more widely around the predicted optimum.

   If the value of the key indicator for the policy test specified in step five is higher than others in the overall set of tests and close to that estimated by the regression model, it can reasonably be accepted as the optimum. If not, all the additional test results are added to those from step three, and steps four to six are repeated until a satisfactory optimum is reached.

7.5 General Conclusions

Our own results suggest that 24 transport model runs may be sufficient to get close to an optimum NPV, even for quite complex problems. We conclude that by making use of our method it is possible to get closer to an optimum, more quickly than via other approaches. Our approach requires considerable skill in interpreting the regression models and in the design of the data set.

REFERENCES

BONSALL, P.W. (1992). PLUTO: Computer assisted training for transport planners. Proc UTSG Conference.

BONSALL, P.W. (1994). PLUTO: Computer assisted training for transport planners. Proceedings of a Conference on computer aided learning in engineering. University of Sheffield, 1994. (Also to appear in Computers and Education, Summer 1995).

BRISTOW, A.L., FOWKES, A.S., BONSALL, P.W. and MAY, A.D. (1994). The optimisation of integrated urban transport strategies: tests based on Edinburgh. Working Paper 425, Institute for Transport Studies, University of Leeds.

GLIM (1992) Royal Statistical Society

APPENDIX 1 - ROAD PROJECT MODEL PARAMETERS

	Intuitive Model (equation 4)		Systematic Model (equation 1)	
Variable	Coefficient	Standard error	Coefficient	Standard error
P1	-6.63	4.2	-14.03	1.5
P2	-10.31	5.2	-8.30	1.5
P3	-18.09	5.2	-12.42	1.4
P4	-	-	-7.99	1.5
P5	20.19	4.5	23.6	1.4
P6	-	-	-	-
P7	-12.71	4.2	-5.03	1.5
P8	-	-	-	-
P9	82.53	5.2	89.79	1.5
P10	34.24	5.2	41.72	1.4
P810			-22.48	4.1
P910			-16.36	3.0
P17			-5.61	4.3
P27			-12.19	4.2
P45			-12.20	4.4
P59			-11.22	2.8
P47			13.75	3.3
P79			9.72	3.0
P13			11.25	4.5
P210			-5.81	4.3
P5710			10.49	4.1
SUB	-18.18	4.2	-	-

Each variable is a dummy

P1 to P10 indicate single projects, 1 to 10.

P810 to P210 indicate an interaction term between the pair of projects indicated in the variable name.

P5710 is a three way interaction term between projects 5, 7 and 10.

SUB is 1 if any two projects directly affect the same road space.

APPENDIX 2 CORDON PRICING

SECTION 1 - ESTIMATION OF OPTIMUM FEE AND NPV PER ZONE USING THE INITIAL MODEL (Equation 5)

Variable	Coefficient	Standard error
FZ1	13.87	4.44
FZ2	23.45	3.03
FZ3	36.24	5.56
FZ4	45.77	3.21
FZ5	31.45	0.98
FZ6	30.32	1.61
$FZ1^2$	-2.49	1.35
$FZ2^2$	-2.96	0.77
$FZ3^2$	-5.30	1.59
$FZ4^2$	-5.86	0.82
$FZ5^2$	-2.77	0.14
$FZ6^2$	-2.57	0.19

Optimum Fees

$$\frac{\partial NPV}{\partial FZ1} = 13.87 - 4.98\ FZ1 = 0 \Rightarrow FZ1 = \frac{13.87}{4.98} = 2.78$$

$$\frac{\partial NPV}{\partial FZ2} = 23.45 - 5.92\ FZ2 = 0 \Rightarrow FZ2 = \frac{23.45}{5.92} = 3.96$$

$$\frac{\partial NPV}{\partial FZ3} = 36.24 - 10.6\ FZ3 = 0 \Rightarrow FZ3 = \frac{36.24}{10.6} = 3.42$$

$$\frac{\partial NPV}{\partial FZ4} = 45.77 - 11.72\ FZ4 = 0 \Rightarrow FZ4 = \frac{45.77}{11.72} = 3.90$$

$$\frac{\partial NPV}{\partial FZ5} = 31.45 - 5.54\ FZ5 = 0 \Rightarrow FZ5 = \frac{31.45}{5.54} = 5.50$$

$$\frac{\partial NPV}{\partial FZ6} = 30.32 - 5.14\ FZ6 = 0 \Rightarrow FZ6 = \frac{30.32}{5.14} = 5.89$$

If FZ1 = 2.8 NPV= 19.3 eg NPV= 13.87 (2.8) - 2.49 $(2.8)^2$

FZ2 = 4.0 NPV= 46.4

FZ3 = 3.4 NPV= 61.9

FZ4 = 3.9 NPV= 89.1

FZ5 = 5.5 NPV= 89.2

FZ6 = 5.9 NPV= 89.4

SECTION 2 - CORDON PRICING MODEL, SYSTEMATIC DATA SET (Equation 6)

Variable	Coefficient	S.E.
FZ1	13.82	2.496
FZ2	26.18	2.625
FZ3	36.30	2.181
FZ4	45.13	2.181
FZ5	37.35	1.636
FZ6	35.58	1.744
$FZ1^2$	-1.976	0.5725
$FZ2^2$	-3.597	0.5926
$FZ3^2$	-5.344	0.4032
$FZ4^2$	-5.398	0.4032
$FZ5^2$	-3.734	0.2268
$FZ6^2$	-3.702	0.2867

Model 6 - 18 run data set

$$\frac{\partial NPV}{\partial FZ1} = 13.82 - 3.95\ FZ_1 = 0 \Rightarrow FZ_1 = \frac{13.82}{3.95} = 3.5$$

$$\frac{\partial NPV}{\partial FZ2} = 26.2 - 7.19\ FZ_2 = 0 \Rightarrow FZ_2 = \frac{26.2}{7.19} = 3.6$$

$$\frac{\partial NPV}{\partial FZ3} = 36.3 - 10.7\ FZ_3 = 0 \Rightarrow FZ_3 = \frac{36.3}{10.7} = 3.4$$

$$\frac{\partial NPV}{\partial FZ4} = 45.1 - 10.8\ FZ_4 = 0 \Rightarrow FZ_4 = \frac{45.1}{10.8} = 4.2$$

$$\frac{\partial NPV}{\partial FZ5} = 37.3 - 7.4\ FZ_5 = 0 \Rightarrow FZ_5 = \frac{37.3}{7.4} = 5.0$$

$$\frac{\partial NPV}{\partial FZ6} = 35.6 - 7.4\ FZ_6 = 0 \Rightarrow FZ_6 = \frac{35.6}{7.4} = 4.8$$

$\therefore$ If FZ1 = 3.5 $\quad NPV = (13.82 \times 3.5) - 1.98\ (3.5^2) = 24.1$

If FZ2 = 3.6, $\quad NPV = (26.2 \times 3.6) - 3.6\ (3.5^2) = 50.2$

If FZ3 = 3.4, $\quad NPV = (36.3 \times 3.4) - 5.3\ (3.4^2) = 62.1$

If FZ4 = 4.2, $\quad NPV = (45.1 \times 4.2) - 5.4\ (4.2^2) = 97.1$

If FZ5 = 5.0, $\quad NPV = (37.3 \times 5) - 3.7\ (5^2) = 94.0$

If FZ6 = 4.8, $\quad NPV = (35.6 \times 4.8) - 3.7\ (4.8^2) = 85.5$

APPENDIX 3

RULE FOR TESTING SIMPLE CURVED FUNCTIONS FOR SUCCESSIVE ZONES GIVEN STARTING VALUE PROVIDED BY AN EXPERT

- With Worked Example

1. Suppose we wish to find the optimum fee to charge at an optimally placed cordon. Suppose that the cordon may be placed around any zone, Z_1 ... Z_n and any fee, f_z, can be changed in each zone z. Let the objective function by Y, which in practice might be NPV.

2. First ask an expert for a rough idea, X^E, of the fee which is likely to be optimum.

3. Secondly, prompt the expert for some the idea of whether this would be expected to get higher(H) or lower(L) as we move out from the centre, or neither.

4. Start with the innermost zone Z_1. Form X^* such that $X^* \approx X^E$, being $< X^E$ if the answer to 3 is H, and $> X^E$ if the answer to 3 is L. Round numbers will be quite adequate and simpler to handle.

5. Perform model runs with a cordon around zone Z_1 and fee $X = X^*$ and $X = 2X^*$. If $Y(2X^*) > Y(X^*)$ then test $X = 4X^*$ otherwise test $X = ½X^*$. This ensures that the optimum is enclosed.

6. Compare the objective function values found, ie Y(X) for these three X values. If the Y's are monotonically increasing we must run $Y(4X^*)$, $Y(8X^*)$, $Y(16X^*)$ etc until one turns down. If the are monotonically decreasing, we must run $Y(X^*/8)$, $Y(X^*/16)$ etc until one these turns down. Then go to step 7.

7. We now move to the next zone. Form new X^* by setting X equal to the fee which maximised the objective function(s) in the previous zone (ie at step 6). Then proceed as as steps 5 and 6 for this new zone.

8. Repeat until all zones have been covered.

9. Enter this data into the regression model, where the 'NPV' is Y (the objective function) there is no constant term, and the fitted exploratory variables are dummy interactions
   ZONE * FEE
   and ZONE * FEE * FEE
   where ZONE is zero/one depending if the observation is for that zone, and FEE is the X value used on that run to get that Y value.

10. Determine the optimum fee for each zone, i, as follows

$$X_i = \frac{-\beta_{1i}}{2\beta_{2i}}$$

where β_{1i} is the coefficient of ZONE * FEE for zone i and β_{2i} is the coefficient of ZONE * FEE * FEE for zone i.

11. Calculate the predicted maximum Y_i for each zone, i, from the estimated equation, using fee level X_i

    ie $\quad Y_i = \beta_{1i} X_i + \beta_{2i} X^2_i$

12. From the Y_i values calculated at step 11, choose the highest. This will give the best zone for the cordon, as well as indicating roughly the optimal fee to be charged. The estimate of objective function Y, say NPV, will only be very approximate, but the choice of zone should be reliable and the optimal fee should not be too far out, particularly considering that it will need to be rounded before any implementation.

APPENDIX 4 BUS FARES AND SERVICE LEVELS (Equation 7)

Variable	Variable Value *	Coefficient	Standard Error
Constant		-97.75	16.8
PKF2	18	-4.27	25.0
PKF3	39	27.44	18.8
PKF4	50	44.71	22.0
PKF5	57	50.20	19.1
PKF6	70	55.03	25.0
PKF7	78	62.53	18.4
PKF8	90	55.16	22.0
OPKF	10	17.57	9.5
OPKF3	15	8.81	10.6
OPKF4	18	3.57	11.2
OPKF5	30	-34.73	11.6
OPKF6	39	-41.41	9.4
OPKF7	79	-334.9	18.5
OPSS		-0.08	0.004
OPS		8.51	0.35

Dummies for flat fares and peak service level perform poorly and so were excluded.

* The actual fares attached to the factor levels. In each case level 1 of the factor is a fare of zero.

Optimal off-peak service level:

$$\frac{\partial NPV}{\partial OPS} = 8.5 - 0.16\ OPS$$

therefore $OPS = 8.5/0.16 = \underline{53}$ roughly

APPENDIX 5: COMBINED TESTS - OPTIMUM INPUT LEVELS

Equation (9), predicted inputs:-

Cordon charge, zone 4 is the optimum location as it has the the highest coefficient and a common squared term (FEES) is used.

$$\frac{\partial NPV}{\partial FZ4} = 72.01 - 14.128\ FEES = 0 \quad \Rightarrow FZ4 = \frac{72.01}{14.128} = \underline{5.10}$$

Peak Fare

$$\frac{\partial NPV}{\partial APF} = 1.517 - 0.017298\ APF = 0 \quad \Rightarrow APF = \frac{1.517}{0.017298} = \underline{88}$$

Off-peak Fare

$$\frac{\partial NPV}{\partial AOF} = 1.033 - 0.12816\ AOF = 0 \quad \Rightarrow AOF = \frac{1.033}{0.12816} = \underline{8}$$

Additional off-peak vehicles

$$\frac{\partial NPV}{\partial AOP} = 1.605 - 0.010296\ AOP = 0 \quad \Rightarrow AOP = \frac{1.605}{0.010296} = \underline{156}$$

Predicted NPV =

$$-37.21 + 1.517\ (88) + 1.033\ (8) - 0.008649\ (88^2) - 0.06408\ (8^2)$$
$$+72.01\ (5.1) - 7.064\ (5.1^2) + 1.605\ (156) - 0.005148\ (156^2)$$

$$= -37.21 + 133.5 + 8.3 - 67.0 - 4.1 + 367.3 - 183.7 + 250.4 - 125.3$$

$$= \underline{£342.2m}$$

K:\...\ALB\REPORTS\WP424