



## **Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin**

### **Copyright statement**

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

### **Liability statement**

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

### **Access Agreement**

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

**Use of High Performance  
Computing in the assessment of  
Congestion Charging in the Greater  
Dublin Area**

**Volume I of II**

**Eoin Alexander O’Cearbhaill**

**Thesis submitted to the University of Dublin,  
Trinity College, for the degree of  
Doctor of Philosophy**

**2004**

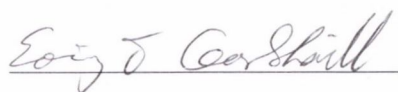
TRINITY COLLEGE  
14 OCT 2004  
LIBRARY DUBLIN

*THOSIS  
7466.1*

## ***Declaration***

The author hereby declares that this thesis, in whole or in part, has not been submitted to any other university as an exercise for a degree. Except where reference has been given in the text, it is entirely the author's own work.

The author confirms that the Library may lend or copy this thesis upon request, for academic purposes.

A handwritten signature in cursive script, reading "Eoin A. O'Cearbhaill", written over a horizontal line.

Eoin A. O'Cearbhaill

June 2004

## *Acknowledgements*

Firstly, I would like to thank my supervisor, Dr. Margaret O'Mahony, who has helped me to develop the skills necessary to complete this research. I also have a special thank you in honour of the late Professor Simon Perry, without whom I would never have made it through my undergraduate degree and without whom I would never have considered a research post-graduate. Dr. Roger West must also be thanked for his help and encouragement during my studies, his advice has always been astute.

Thanks also to the fellow post graduate students, in particular those I have shared an office with over the years including Stephen, Quentin, Warren, Clare, Markus and Brian.

I would also like to thank my family for all there love and support throughout the course of my studies. They have been a huge source of help, support and encouragement. In fact there are numerous people, too many to mention here, be they friends or loved ones, who have helped me and encouraged me along the way and I am most grateful to them for their time, patience, and support and encouragement.

Finally, I would like to thank the Irish Association for High Performance Computing in Ireland for their funding of this research.

## *Summary*

Traffic demand management is part of the package of measures currently being implemented to offset the imbalance between transport demand and supply on the transportation network of the Greater Dublin Area (GDA). An analysis of the development of the transport environment in the GDA leads to the conclusion that the transportation deficit that exists in the GDA can only be addressed using an integrated transportation strategy involving infrastructural and service improvements on one side and a demand management strategy on the other. This transportation deficit has resulted from delays in infrastructure projects and unrealistic planning targets due to institutional and regulatory inefficiencies, cumbersome and time consuming planning system, and, the high economic growth since the Dublin Transportation Initiative began in 1988.

The Dublin Transportation Network Model (DTNM), which uses the SATURN (Simulation and Assignment of Traffic to Urban Road Networks) traffic assignment and simulation model, is used to analyse potential solutions over the GDA. The SATURN model is assessed with respect to the DTNM, with a view to finding the areas within the model that take up the most Central Processing Unit (CPU) time, and to assess the potential for reprogramming these areas to run in parallel. The process of reprogramming the most time consuming, or computationally intensive, areas of the model in parallel is then undertaken in order to produce a more efficient model that can then be used to assess the effects of road user charging schemes on the DNM. This process results in a more efficient use of computer time over a variety of network configurations and sizes and an improved efficiency in examining the DTNM.

Results from this research include outputs from the parallel model that demonstrate its ability to converge while retaining at least the same accuracy as that of the sequential model. The performance of the new parallel model is also compared with that of the sequential (standard format) model to demonstrate a performance enhancement. An increase in performance, measured using the variable 'speed-up', of 8.16 for a 16-processor parallel implementation on the IBM supercomputing platform is achieved.

Further to this, the impact of a series of inner city cordon congestion charging scenarios designed to reduce the demand for travel on the GDA highway network is evaluated. The effects of this cordon application should help to counteract the transportation deficit on the highway network. In order to examine whether this is the case or not, the newly developed parallelised version of the traffic network analysis model, running across parallel processors on the SP2 super computing platform, is utilised to run test scenarios of the inner city cordon congestion charging scheme using a number of different charge levels and a range of different price elasticities of user demand, which are based on the conclusions reached from a comprehensive literature review of international experience and academic research.

The analysis of this inner city congestion cordon scheme evaluates the impacts of inserting an inner city cordon onto the highway network. The responses of trip makers whose class of vehicle is either heavy goods vehicle (HGVs) or private car is assessed for a range of charge levels and elasticity estimates. Variables used in the examination include travel time, travel distance, queue lengths and average speed.

Conclusions are drawn from the output of this modelling evaluation of an inner city congestion charging scheme and its affects on the GDA. Overall, it is found that congestion in the GDA reduces by a significant amount as a result of the inner city cordon. There is an approximate reduction of 14% in over-capacity queuing and a 15 to 16% reduction in the total travelled time on the highway network for an elasticity of – 0.3 and a charge of €5.70.

TABLE OF CONTENTS

SUMMARY I

**TABLE OF CONTENTS** ..... **III**

**LIST OF TABLES** ..... **IX**

**LIST OF FIGURES** ..... **X**

**LIST OF EQUATIONS** ..... **XII**

**ABBREVIATIONS** ..... **XIII**

**1. INTRODUCTION** ..... **1**

1.1 BACKGROUND ..... 2

1.2 AIM ..... 3

1.3 RESEARCH APPROACH ..... 4

1.4 OUTLINE OF THESIS ..... 5

**2. DEVELOPMENT OF THE TRANSPORTATION ENVIRONMENT IN THE GREATER DUBLIN AREA** ..... **8**

2.1 INTRODUCTION ..... 9

2.2 TRANSPORT PLANNING IN THE GREATER DUBLIN AREA ..... 10

2.2.1 INTRODUCTION ..... 10

2.2.2 THE IRISH PLANNING SYSTEM, 1963 – 1999 ..... 12

2.2.3 THE IRISH PLANNING SYSTEM, 1999 – ..... 14

2.2.4 EXISTING INSTITUTIONAL ARRANGEMENTS FOR TRANSPORT AND LAND USE IN THE GDA ..... 17

2.2.5 PROPOSED INSTITUTIONAL ARRANGEMENTS FOR TRANSPORT AND LAND USE IN THE GDA ..... 20

2.3 TRANSPORTATION DEVELOPMENT IN THE GDA PRIOR TO THE DUBLIN TRANSPORTATION INITIATIVE (DTI) ..... 22

2.3.1 INTRODUCTION ..... 22

2.3.2 THE MYLES WRIGHT REPORT (1967) ..... 23

2.3.3 SUBSEQUENT STUDIES ..... 25

2.4 THE DUBLIN TRANSPORTATION INITIATIVE (DTI) ..... 26

2.4.1 INTRODUCTION ..... 26

2.4.2 DTI METHODOLOGY ..... 28

2.4.3 RECOMMENDED STRATEGY OF THE DTI ..... 33

2.4.4 TRANSPORTATION MODELLING AND THE DTI ..... 36

2.4.4.1 THE DTI MODEL APPROACH ..... 36

2.4.4.2 DATA COLLECTION ..... 37

2.4.4.3 DTI TRIP GENERATION AND FUTURE YEAR MATRICES ..... 39

2.4.4.4 GENERALISED COST AND MODE SPLIT ..... 40

2.4.4.5 ASSIGNMENT ..... 42



2.4.5	IMPLEMENTATION OF THE DTI MEASURES .....	43
2.5	DUBLIN'S DEVELOPMENT SINCE THE DTI BEGAN .....	46
2.5.1	ECONOMIC GROWTH AND POPULATION CHANGE .....	46
2.5.2	INCREASED TRAVEL DEMAND .....	49
2.5.3	DUBLIN'S TRANSPORTATION DEFICIT .....	52
2.6	FUTURE DEVELOPMENT IN THE GREATER DUBLIN AREA .....	53
2.6.1	THE STRATEGIC PLANNING GUIDELINES FOR THE GREATER DUBLIN AREA .....	53
2.6.2	AN INTEGRATED TRANSPORTATION STRATEGY FOR THE GREATER DUBLIN AREA .....	57
2.6.2.1	<i>INFRASTRUCTURE AND SERVICE IMPROVEMENTS</i> .....	59
2.6.2.2	<i>DEMAND MANAGEMENT</i> .....	61
2.6.3	LIKELY BENEFITS TO THE GREATER DUBLIN AREA .....	62
2.7	SUMMARY OF THE DEVELOPMENT WITHIN THE GREATER DUBLIN AREA .....	65
<b>3</b>	<b>ROAD USER CHARGING AND THE GREATER DUBLIN AREA: LITERATURE REVIEW.....</b>	<b>69</b>
3.1	INTRODUCTION .....	70
3.2	DEVELOPMENT OF ROAD USER CHARGING .....	71
3.2.1	INTRODUCTION .....	71
3.2.2	TRAFFIC FLOW AND CONGESTION ECONOMICS .....	73
3.2.3	CONGESTION CHARGING.....	76
3.3	URBAN ROAD USER CHARGING AND IMPLEMENTATION.....	80
3.3.1	INTRODUCTION .....	80
3.3.2	BASIC DESIGN CRITERIA.....	81
3.3.3	PRACTICAL ASPECTS OF DESIGN AND IMPLEMENTATION .....	82
3.4	POLITICAL AND SOCIAL ACCEPTABILITY OF ROAD USER CHARGING .....	85
3.4.1	INTRODUCTION .....	85
3.4.2	ACCEPTABILITY OF ROAD USER CHARGING.....	86
3.4.3	INCREASING ACCEPTABILITY OF ROAD USER CHARGING.....	88
3.5	EFFECTIVENESS OF ROAD USER CHARGING .....	92
3.5.1	INTRODUCTION .....	92
3.5.2	STUDIED EFFECTS OF ROAD USER CHARGING .....	92
3.6	ROAD USER CHARGING STUDIES; PRE AND POST IMPLEMENTATION .....	99
3.6.1	INTRODUCTION .....	99
3.6.2	ROAD USER CHARGING OF A CITY CENTRE .....	100
3.6.2.1	<i>ROAD USER CHARGING IN THE GREATER LONDON AREA</i> .....	100
3.6.2.2	<i>ROAD USER CHARGING IN SINGAPORE</i> .....	102
3.6.2.3	<i>HONG KONG'S ELECTRONIC ROAD PRICING TRIAL</i> .....	103

3.6.2.4	CONGESTION SPECIFIC PRICING FOR CAMBRIDGE, ENGLAND .....	105
3.6.3	CITY CENTRE TOLL RINGS; SCANDINAVIAN EXPERIENCE.....	106
3.6.3.1	NORWAY'S URBAN TOLL RINGS.....	106
3.6.3.2	ROAD USER CHARGING IN STOCKHOLM .....	108
3.6.4	AREA-WIDE ROAD USER PRICING; THE RANDSTAD REGION, THE NETHERLANDS .....	109
3.6.5	SINGLE FACILITY ROAD USER PRICING: CALIFORNIA'S PRIVATE TOLL LANES .....	111
3.7	PRICE ELASTICITY OF USER DEMAND .....	111
3.7.1	INTRODUCTION .....	111
3.7.2	REVIEW OF PRICING ELASTICITIES.....	112
3.8	CONCLUSIONS .....	116
3.8.1	EFFECTIVENESS OF ROAD USER CHARGING .....	116
3.8.2	POLITICAL AND SOCIAL ACCEPTABILITY OF ROAD USER PRICING .....	116
3.8.3	ROAD USER CHARGING STUDIES; PRE AND POST IMPLEMENTATION .....	117
3.8.4	PRICE ELASTICITIES OF USER DEMAND .....	118
<b>4</b>	<b>IMPROVING THE PERFORMANCE OF THE DUBLIN TRANSPORTATION NETWORK MODEL USING HIGH PERFORMANCE COMPUTING TECHNIQUES.....</b>	<b>119</b>
4.1	INTRODUCTION .....	120
4.2	THE DUBLIN TRANSPORTATION OFFICE TRAFFIC MODEL.....	123
4.2.1	INTRODUCTION .....	123
4.2.2	ZONING AREAS OF INTEREST .....	124
4.2.3	TRANSPORTATION NETWORKS.....	125
4.2.4	THE TRANSPORTATION MODEL STRUCTURE: CLASSIC 40STAGE MODEL.....	127
4.2.4.1	TRIP GENERATION - ATTRACTION STAGE.....	128
4.2.4.2	TRIP DISTRIBUTION STAGE .....	129
4.2.4.3	MODAL SPLIT STAGE .....	129
4.2.4.4	ASSIGNMENT STAGE .....	130
4.2.5	OVERVIEW OF THE DTO TRAFFIC MODEL LAYOUT .....	132
4.3	THE SATURN TRAFFIC SIMULATION AND ASSIGNMENT MODEL .....	134
4.3.1	INTRODUCTION .....	134
4.3.2	THE SATURN MODEL FRAMEWORK.....	135
4.4	INITIAL ANALYSIS OF SATURN .....	138
4.4.1	INTRODUCTION .....	138
4.4.2	INITIAL ANALYSIS.....	139
4.4.3	CONCLUSION .....	141
4.5	TRIP ASSIGNMENT .....	143

4.5.1	INTRODUCTION .....	143
4.5.2	MULTIPLE USER CLASSES .....	144
4.5.3	USER EQUILIBRIUM ASSIGNMENT .....	144
4.5.4	THE FRANK-WOLFE ALGORITHM .....	146
4.5.5	PARTAN ASSIGNMENT .....	148
4.5.6	METHOD OF SUCCESSIVE AVERAGES (MSA) .....	148
4.5.7	CONCLUSION .....	149
4.6	SHORTEST PATHS .....	150
4.6.1	INTRODUCTION .....	150
4.6.2	TREE BUILDING ON TRANSPORTATION NETWORKS .....	151
4.6.3	SHORTEST PATH ALGORITHMS FOR TRANSPORTATION NETWORKS .....	153
4.6.4	CONCLUSIONS .....	159
4.7	REVIEW OF PARALLEL PROGRAMMING IN THE AREA OF TRAFFIC ASSIGNMENT AND SHORTEST PATH CALCULATION .....	160
4.7.1	INTRODUCTION .....	160
4.7.2	PARALLEL COMPUTING ARCHITECTURE .....	162
4.7.3	PARALLEL COMPUTATION: METHODS AND CONSIDERATIONS .....	164
4.7.3.1	<i>DECOMPOSITION</i> .....	164
4.7.3.2	<i>COMPUTATIONAL AND STORAGE DEPENDENCE</i> .....	167
4.7.3.3	<i>INTER-PROCESSOR COMMUNICATION</i> .....	167
4.7.4	PERFORMANCE ANALYSIS OF PARALLEL SYSTEMS .....	170
4.7.4.1	<i>INTRODUCTION</i> .....	170
4.7.4.2	<i>SPEED-UP, RELATIVE BURDEN AND EFFECIENCY</i> .....	171
4.7.4.3	<i>THE AVERAGE PARALLELISM</i> .....	173
4.7.4.4	<i>BENEFIT VERSES COST PROFILES</i> .....	174
4.7.5	PARALLEL PROCESSING IN TRAFFIC ASSIGNMENT AND SHORTEST PATH THEORY .....	175
4.7.5.1	<i>ASSIGNMENT</i> .....	175
4.7.5.2	<i>SINGLE SOURCE ONE-TO-ALL SHORTEST PATH PARALLEL ALGORITHM</i> .....	178
4.7.6	CONCLUSION .....	182
<b>5</b>	<b>PARALLEL IMPLEMENTATION OF A TRANSPORT NETWORK MODEL</b> .....	<b>183</b>
5.1	INTRODUCTION .....	184
5.2	METHOD AND STRATEGY FOR THE PARALLELISATION OF THE SEQUENTIAL SATURN MODEL .....	185
5.2.1	INTRODUCTION .....	185
5.2.2	BACKGROUND .....	186
5.2.3	THE IBM SP2 SUPERCOMPUTER .....	188
5.2.4	PARALLELISATION STRATEGY .....	191

5.2.5	PROGRAMMING THE SINGLE-SOURCE ONE-TO-ALL SHORTEST PATH ALGORITHM IN PARALLEL.....	193
5.3	RESULTS AND DISCUSSION.....	200
5.3.1	INTRODUCTION.....	200
5.3.2	CONVERGENCE.....	202
5.3.3	COMPARISON OF MODEL OUTPUTS.....	204
5.3.4	PERFORMANCE.....	205
<b>6</b>	<b>INNER CITY CONGESTION CHARGING IN THE GREATER DUBLIN AREA .....</b>	<b>213</b>
6.1	INTRODUCTION.....	214
6.2	BACKGROUND.....	216
6.2.1	TRANSPORTATION DEFICIT.....	216
6.2.2	ROAD USER CHARGING.....	219
6.2.3	INNER CITY CONGESTION CHARGING CORDON SCHEME.....	220
6.3	METHODOLOGY.....	222
6.3.1	INTRODUCTION.....	222
6.3.2	PREVIOUS STUDIES OF ROAD USER CHARGING IN THE GDA.....	224
6.3.3	CORDON MODELLING INPUT DATA.....	225
6.3.3.1	<i>MODEL INPUT DATA: THE HIGHWAY NETWORK.....</i>	<i>225</i>
6.3.3.2	<i>MODEL INPUT: THE TRIP MATRIX.....</i>	<i>228</i>
6.3.4	CHARGE LEVELS AND THE ELASTICITY OF USER RESPONSE.....	229
6.3.4.1	<i>CHARGE LEVELS.....</i>	<i>230</i>
6.3.4.2	<i>PRICE ELASTICITY OF USER DEMAND.....</i>	<i>230</i>
6.3.5	ELASTICITY DEMAND FUNCTION.....	231
6.3.6	GENERALISED COST.....	233
6.3.7	INNER CITY CONGESTION CHARGING CORDON TESTING.....	234
6.3.8	TEST SCENARIOS.....	236
6.3.8.1	<i>LIST OF TESTS: 2006.....</i>	<i>237</i>
6.3.8.2	<i>LIST OF TESTS: 2000.....</i>	<i>237</i>
6.3.8.3	<i>LIST OF TESTS: 1997.....</i>	<i>238</i>
6.4	RESULTS AND DISCUSSION.....	239
6.4.1	INTRODUCTION.....	239
6.4.2	DISTANCES TRAVELLED OVER THE HIGHWAY NETWORK.....	240
6.4.3	TRAVEL TIME.....	245
6.4.4	SPEED OVER THE HIGHWAY NETWORK.....	249
6.4.5	QUEUING.....	252
6.4.6	FUEL CONSUMPTION.....	255
<b>7</b>	<b>CONCLUSIONS.....</b>	<b>258</b>

7.1	CONCLUSIONS FROM PARALLEL IMPLEMENTATION .....	259
7.2	CONCLUSIONS FROM INNER CITY CONGESTION CHARGING IMPLEMENTATION.....	260
	<b>REFERENCES</b> .....	<b>264</b>
	<b>APPENDIX 1: DTI OBJECTIVES/CRITERIA AND SUB-CRITERIA</b> .....	<b>289</b>
	<b>APPENDIX 2: LITTLE ENDIAN AND BIG ENDIAN</b> .....	<b>292</b>
	<b>APPENDIX 3: FORTRAN CODE DESCRIPTION OF THE ASSIGNMENT AND SHORTEST PATH ALGORITHMS WITHIN THE SATURN MODEL</b> .....	<b>294</b>
	<b>APPENDIX 4: MPI IMPLEMENTATION</b> .....	<b>323</b>
	<b>APPENDIX 5: MPI SAMPLE CODING TAKEN FROM AN ASSIGNMENT ROUTINE WITHIN SATALL</b> .....	<b>328</b>
	<b>APPENDIX 6: MODEL HARFWARE</b> .....	<b>360</b>
	<b>APPENDIX 7: RESULTS FROM 1997 CONGESTION CHARING TESTS</b> .....	<b>363</b>
	<b>APPENDIX 8 RESULTS FROM 2000 CONGESTION CHARING TESTS</b> .....	<b>380</b>
	<b>APPENDIX 9 RESULTS FROM 2006 CONGESTION CHARING TESTS</b> .....	<b>416</b>
	<b>APPENDIX 10: SAMPLE NETWORK FILE REPRESENTING THE 2000 DUBLIN TRANSPORTATION NETWORK WITH INNER CITY CORDON APPLICATION</b> .....	<b>460</b>
	<b>APPENDIX 11: DTO ZONES</b> .....	<b>477</b>
	<b>APPENDIX 12: SAMPLE MATRIX FILE REPRESENTING THE 2006 TRIP DEMAND MATRIX</b> .....	<b>493</b>

## *List of Tables*

TABLE 2.1: GROWTH IN DUBLIN BETWEEN 1991 AND 2011.....	31
TABLE 2.2: PERFORMANCE OF SUB-CRITERIA IN DTI.....	32
TABLE 2.3: ESTIMATED FINISHING DATES FOR DTI STRATEGY MEASURES .....	44
TABLE 2.4: FACTORS AFFECTING TRANSPORT DEMAND, 1998 .....	46
TABLE 2.5: FACTORS AFFECTING TRANSPORT DEMAND, 2000 .....	47
TABLE 2.6: TRANSPORT TRIP DEMAND AND GROWTH, 1998 .....	50
TABLE 2.7: TRANSPORT TRIP DEMAND AND GROWTH, 2000 .....	51
TABLE 2.8: PRIVATE CARS LICENSED FOR THE FIRST TIME .....	52
TABLE 2.9: ESRI ALTERNATIVE GROWTH PREDICTIONS .....	55
TABLE 2.10: POPULATION AND HOUSEHOLD PREDICTIONS .....	56
TABLE 2.11: AVERAGE 2016 AM PEAK PERIOD JOURNEY TIMES.....	63
TABLE 2.12: PREDICTED TRANSPORTATION MODE SPLIT IN THE MORNING PEAK.....	64
TABLE 3.1: IMPACTED BEHAVIOURAL GROUPS.....	79
TABLE 3.2: INTEREST GROUPS .....	87
TABLE 3.3: ACCEPTABILITY OF ROAD USER PRICING .....	87
TABLE 3.4: EFFECTIVENESS OF ROAD USER CHARGING .....	94
TABLE 4.1: PROFILE OF CPU RELATED FIGURES .....	142
TABLE 4.2: NETWORK DIMENSIONS AND CPU TIMES.....	158
TABLE 5.1: SEQUENTIAL AND PARALLEL OUTPUT COMPARED .....	204
TABLE 5.2: PERFORMANCE PARAMETERS FOR THE PARALLEL IMPLEMENTATION.....	207
TABLE 5.3: EXECUTION TIMES FOR THE PARALLEL MODEL .....	211
TABLE 6.1: FACTORS INFLUENCING TRAFFIC GROWTH .....	218
TABLE 6.2: 2006 TEST FILES.....	237
TABLE 6.3: 2000 TEST FILES.....	238
TABLE 6.4: 1997 TEST FILES.....	239
TABLE 6.5: DISTANCES TRAVELLED ON BUFFER CENTROID CONNECTORS.....	245
TABLE 6.6: FLOW COEFFICIENTS FOR PCU MEASUREMENT .....	247
TABLE 6.7: FUEL CONSUMPTION STATISTICS.....	257

## *List of Figures*

FIGURE 1.1: THESIS OUTLINE .....	6
FIGURE 2.1: BASIC MAP OF THE GREATER DUBLIN AREA.....	9
FIGURE 2.2: STRATEGIC MAP OF THE GREATER DUBLIN AREA.....	11
FIGURE 2.3: COHESIVE PLANNING SYSTEM .....	15
FIGURE 2.4: EXISTING TRANSPORTATION INSTITUTIONAL ARRANGEMENTS .....	18
FIGURE 2.5: PROPOSED TRANSPORTATION INSTITUTIONAL ARRANGEMENTS .....	20
FIGURE 2.6: THE VISION OF DUBLIN.....	28
FIGURE 2.7: DTI THEME BASED APPROACH.....	29
FIGURE 2.8: DTI LAND USE SCENARIOS.....	30
FIGURE 2.9: POPULATION STATISTICS .....	48
FIGURE 2.10: STRATEGIC ROAD MAP .....	59
FIGURE 2.11: PUBLIC TRANSPORT NETWORKS.....	60
FIGURE 3.1: VEHICLE TRAJECTORIES.....	74
FIGURE 3.2: COST VERSUS VOLUME.....	74
FIGURE 3.3: ECONOMIC COMPONENTS INVOLVED IN ROAD USER PRICING.....	78
FIGURE 3.4: ACCEPTABILITY OF DEMAND MANAGEMENT MEASURES.....	90
FIGURE 3.5: GOODWIN'S RULE OF THREE.....	91
FIGURE 4.1: OUTLINE OF THE AREAS ASSESSED IN CHAPTER 4 .....	122
FIGURE 4.2: A TYPICAL ZONING SYSTEM.....	125
FIGURE 4.3: TRANSPORT NETWORK NODES AND LINKS .....	126
FIGURE 4.4: TRANSPORT ZONES, CENTROIDS, NODES AND LINKS .....	127
FIGURE 4.5: FOUR-STAGE MODEL .....	128
FIGURE 4.6: SPPED-FLOW AND COST-FLOW CURVES.....	131
FIGURE 4.7: OVERVIEW OF THE DTO TRAFFIC MODEL .....	133
FIGURE 4.8: SATURN SIMULATION AND ASSIGNMENT MODEL.....	135
FIGURE 4.9: THE THEORETICAL OPTIMUM SPEED-UP FOR A PARALLEL PROGRAM .....	170
FIGURE 5.1: IBM SP2 SYSTEM PARAMETERS .....	188
FIGURE 5.2: DATA FLOW FOR EACH PARALLEL PROCESSOR .....	196
FIGURE 5.3: INTER-PROCESSOR COMMUNICATION.....	197
FIGURE 5.4: SPEED-UP PROFILES .....	208
FIGURE 5.5: EFFICIENCY PROFILE .....	209
FIGURE 5.6: EFFICIENCY-EXECUTION TIME PROFILE .....	210
FIGURE 6.1: BASIC MAP OF THE GREATER DUBLIN AREA.....	222
FIGURE 6.2: MAP SHOWING CORDON LOCATION.....	223
FIGURE 6.3: MAP OF THE COARSE ZONES IN THE GDA TRANSPORT NETWORK.....	227

FIGURE 6.4: DISTANCE TRAVELLED IN THE SIMULATION PERIOD .....	241
FIGURE 6.5: DISTANCE TRAVELLED OVER THE SIMULATION NETWORK.....	242
FIGURE 6.6: DISTANCE TRAVELLED OVER THE BUFFER NETWORK.....	244
FIGURE 6.7: DISTANCE TRAVELLED OVER THE BUFFER CENTROID CONNECTORS .....	244
FIGURE 6.8: TOTAL TRAVEL TIME IN THE SIMULATION NETWORK .....	247
FIGURE 6.9: TOTAL TRAVEL TIME IN THE BUFFER NETWORK.....	248
FIGURE 6.10: TOTAL TRAVEL TIME IN THE SIMULATION PERIOD.....	249
FIGURE 6.11: EFFECT OF CHARGE AND ELASTICITY ON AVERAGE SPEED .....	250
FIGURE 6.12: AVERAGE SPEED ON THE BUFFER NETWORK.....	251
FIGURE 6.13: AVERAGE SPEED ON THE BUFFER NETWORK EXCLUDING BUFFER CENTROID CONNECTORS.....	252
FIGURE 6.14: QUEUING ON THE 2006 HIGHWAY NETWORK .....	253
FIGURE 6.15: LINK CRUISE TIMES FOR THE SIMULATION NETWORK.....	254
FIGURE 6.16 OVER-CAPACITY QUEUING .....	255



## *List of Equations*

EQUATION 2.1: OUTPUT FUTURE YEAR TRIP MATRIX.....	39
EQUATION 2.2: BALANCING FACTOR.....	40
EQUATION 2.3: BALANCING FACTOR.....	40
EQUATION 2.4: GENERALISED COST FOR PRIVATE CAR.....	40
EQUATION 2.5: GENERALISED COST FOR PUBLIC TRANSPORT .....	41
EQUATION 2.6: HIGHER NESTED LOGIC FORM OF MODE CHOICE MODEL.....	41
EQUATION 2.7: CAR/PUBLIC TRANSPORT NESTED FORM OF MODE CHOICE MODEL .....	42
EQUATION 2.8: GENERALISED COST, LOGSUM AVERAGE.....	42
EQUATION 4.1: TRIP GENERATION EQUATION .....	128
EQUATION 4.2: TRIP ATTRACTION EQUATION .....	128
EQUATION 4.3: TRIP DISTRIBUTION FOR ORIGIN I .....	129
EQUATION 4.4: TRIP DISTRIBUTION FOR ORIGIN J .....	129
EQUATION 4.5: GENERALISED COST OF TRAVEL BY PRIVATE CAR .....	130
EQUATION 4.6: OBJECTIVE FUNCTION FOR A SINGLE USER CLASS .....	145
EQUATION 4.7: OBJECTIVE FUNCTION FOR A MULTIPLE USER CLASS.....	145
EQUATION 4.8: TO GENERATE A SET OF LINK FLOWS IN FRANK-WOLFE .....	147
EQUATION 4.9: TO GENERATE A SET OF LINK FLOWS IN PARTAN.....	148
EQUATION 4.10: TO GENERATE A SET OF LINK FLOWS IN MSA .....	149
EQUATION 4.11: SPEED-UP .....	171
EQUATION 4.12: RELATIVE BURDEN .....	172
EQUATION 6.1: ELASTICITY DEMAND FUNCTION .....	232
EQUATION 6.2: GENERALISED COST FOR PRIVATE CAR.....	233
EQUATION 6.3: GENERALISED COST FOR PUBLIC TRANSPORT .....	233
EQUATION 6.4: FUEL CONSUMPTION EQUATION.....	256

## ***Abbreviations***

ALS – Area Licensing Scheme

AVC – Average Variable Curve

BAC – Bus Atha Cliath

BE – Bus Eireann

CBA – Cost Benefit Analysis

CEC – Commission of the European Communities

CIE – Córas Iompair Éireann

CONTRAM – Continuance Traffic Assignment Mode

CSO – Central Statistic’s Office

CPU – Central Processing Unit

DART – Dublin Area Rapid Transit

DED – District Electoral Division

DEHLG - Department of the Environment, Heritage and Local Government

DoE – Department of the Environment

DoELG – Department of the Environment and Local Government

DoT – Department of Transport

DPE – Department of Public Enterprise

DRRTS – Dublin Rapid Transit Study

DSRC – Dedicated Short Range Communications

DTI – Dublin Transportation Initiative

DTNM – Dublin Transportation Network Model

DTO – Dublin Transportation Office

DTS – Dublin Transport Study

EIS – Environmental Impact Statement

EIA - Environmental Impact Assessment

EPA – Environmental Protection Agency

ERDO – Eastern Regional Development Organisation

ERP – Electronic Road Pricing

ESRI – Economic and Social Research Institute

ETM – Electronic Ticketing Machine

EU – European Union

FDDI – Fibre Distributed Data Interface

FIFO – First-In First-Out

GB - Gigabyte

GDA – Greater Dublin Area

GDA – Greater Dublin Area Authority

GIS – Geographic Information System

GNP – Gross National Product

GRD – Gross Residential Density

HGV – Heavy Goods Vehicle

HLGTIC – High Level Group on Transport Infrastructure Charging

HOT – High Occupancy Toll

HPC – High Performance Computing

IBG – Impact Behavioural Group

IE – Iarnrod Eireann

IFSC – International Financial Services Centre

I/O – Input/Output

ISO – International Organisation for Standardisation

ITS – Institute of Transport Studies

IU – In-vehicle Unit

LAN – Local Area Network

LRPO – Light Rail Project Office

LRT – Light Rail Transit

LT – London Transport

MASL – Multiple Assignment Simulation Loops

MB - Megabyte

MISD – Multiple Instruction, Single Data

MIMI – Multiple Instruction, Multiple Data

MIPS – Millions of Instructions Per Second

MSA – Method of Successive Averages

MPI – Message Passing Interface

NO<sub>x</sub> – Nitrogen Oxide

NPROC – Number of Processors

NRA – National Roads Authority

O-D – Origin – Destination

PARTAN – Parallel Tangent

PTMT – Public Transport Mode Choice

PVM – Parallel Virtual Machine

QBC – Quality Bus Corridor

RAM – Random Access Memory

RISC – Reduced Instruction Set Computer

RPA – Railway Procurement Agency

ROCOL – Road Charging Options for London

SATCHMO – Saturn Travel Choice Model

SATURN – Simulation and Assignment of Traffic in Urban Road Networks

SIMD – Single Instruction, Multiple Data

SISD – Single Instruction, Single Data

SO<sub>2</sub> – Sulphur Dioxide

SPG – Strategic Planning Guidelines

SPGO – Strategic Planning Guidelines Office

SPMD – Single Program, Multiple Data

TCC – Transport Consultative Commission

TORG – Transport Operations Research Group

TRL – Transportation Research Laboratory

UFA – Unformatted Assignment

UFM – Unformatted Matrix

UFN – Unformatted Network

UFS – Unformatted Simulation

UK DETR – UK Department of the Environment, Transport and the Regions

VPS – Vehicle Positioning System

# *Chapter 1*

## *Introduction*

## ***1.1 Background***

Traffic demand management is part of the package of measures currently being implemented to offset the imbalance between transport demand and supply on the transportation network of the Greater Dublin Area (GDA). An analysis of the development of the transport environment in the GDA leads to the conclusion that the transportation deficit that exists in the GDA can only be addressed using an integrated transportation strategy involving infrastructural and service improvements on one side and a demand management strategy on the other. This transportation deficit has resulted from delays in infrastructure projects and unrealistic planning targets due to institutional and regulatory inefficiencies, cumbersome and time consuming planning system, and, the high economic growth since the Dublin Transportation Initiative began in 1988.

In 1999, the Strategic Planning Guidelines (SPG), a “framework for integrated land use and transportation for the sustainable development of the GDA up to the year 2011”, was produced, which provided the first co-ordinated settlement plan for the GDA and the starting point for the latest transportation plans to address traffic congestion (Brady Shipman Martin et al., 1999). From a transportation perspective, the SPG objectives are to consolidate development and increase density in the GDA so as to provide a basis for a better public transport system, and, to facilitate a shift away from the private car to public transport. As such, the SPG recommend that demand management measures should be incorporated into all strategic planning and implementation to reduce traffic demand, dissuade private car commuting and encourage the use of public transport. In turn, the DTO 2000-2016 Strategy also advocated an integrated transportation strategy involving demand management and infrastructural and service improvements. It was also recognised that a comprehensive study is necessary to develop a demand management strategy to reduce growth in overall travel by private car and to effect further modal transfer from private car to public transport modes.

‘A Study of Road Pricing in Dublin’ was produced by Oscar Faber et al in 1998. This is the only report that has been produced to assess road user charging on an area-wide



basis in the GDA, and, in essence was a feasibility study. It examined a small number of testing scenarios to try to assess the feasibility of road user charging and other fiscal instruments, such as parking charges, for the management of transportation and traffic demand in the GDA. The study also looked at the different forms of road user charging and their implementation and went on to confirm the potential of road user charging in contributing to the management of traffic in the GDA. The study was undertaken strictly as a preliminary study to assess the potential of road user charging and it was assumed that further information and detail would be required to evaluate in full the charging schemes in the GDA. Based on the positive preliminary findings of the study the report recommends that further work should be undertaken to investigate the feasibility of introducing a road user charging scheme, and in particular an inner city cordon charging scheme, to the GDA highway network.

### *1.2 Aim*

The first main objective of this research was to assess the SATURN model (van Vliet, 1982), as applied to the Dublin Transportation Network Model (DTNM) with a view to finding the areas within the model that take up the most Central Processing Unit (CPU) time, and to reprogram these areas to run in parallel so that higher levels of efficiency can be achieved (Dublin Transportation Initiative, 1995). Reprogramming the most time consuming, or computationally intensive, areas of the sequential model in parallel should allow for a more efficient use of computer time over a variety of network configurations and sizes, and in doing so, result in improved efficiency in examining the DTNM. The overall objectives for the parallel implementation are to:

- Demonstrate that the modified parallel model converges and produces results that are comparable to the sequential original; and to,
- Evaluate the efficiency of the parallel implementation when more than one processor is used, i.e., 2, 4, 8, and 16 processors.

Secondly, the impact of a series of inner city cordon congestion charging scenarios designed to reduce the demand for travel on the GDA highway network is evaluated.

The effects of this cordon application should help to counteract the transportation deficit on the GDA highway network. In order to examine whether this is the case or not, the newly developed parallelised version of the traffic network analysis model, running across parallel processors on the SP2 super computing platform, is utilised to run test scenarios of the inner city cordon congestion charging scheme using a number of different charge levels and a range of different price elasticities of user demand, which are based on the conclusions reached from a comprehensive literature review of international experience and academic research.

### ***1.3 Research Approach***

A number of approaches were required for this work. Initially a literature-based study was required for the analysis of the development of the GDA. Some information was also gathered for this analysis using observation and through correspondence. A literature-based study was also required for the examination of the theory and background of road user charging and high performance computing techniques and practices.

To gain an understanding of the methodologies used in the SATURN model a number of communiqués were made with the creator of the model, Dr. Dirck van Vliet and with the primary users of the SATURN model in the GDA, the Dublin Transportation Office (DTO). Additionally, a visit to the Institute of Transport Studies (ITS), Leeds University, was organised to meet with Dr. van Vliet. Other sources of information have included communiqués with the Central Statistics Office (2002), and, Origin-Destination (O-D) data from the DTO.

An extensive amount of programming and program testing was required to develop the parallel implementation of the sequential SATURN model and to get the parallel implementation to run on the IBM Supercomputer. A knowledge of a number of operating systems (DOS, UNIX, LINUX), programming languages (Fortran 77, C++, C, MPI) and compilers (FTN77, f77, g77, c90) was necessary to complete the parallel implementation. A one year high performance computing Masters taught course was attended to help in this regard and to gain an understanding of how complex parallel

platforms operate, and, how parallel algorithms are produced to exploit system architecture and other characteristics.

Similarly, once the parallel implementation was complete two extensive testing programs were undertaken to ensure the accuracy, efficiency and performance of the new parallel model, and once this had been achieved, to develop the methodology and produce the cordon models for the different years, charge levels and elasticities tested.

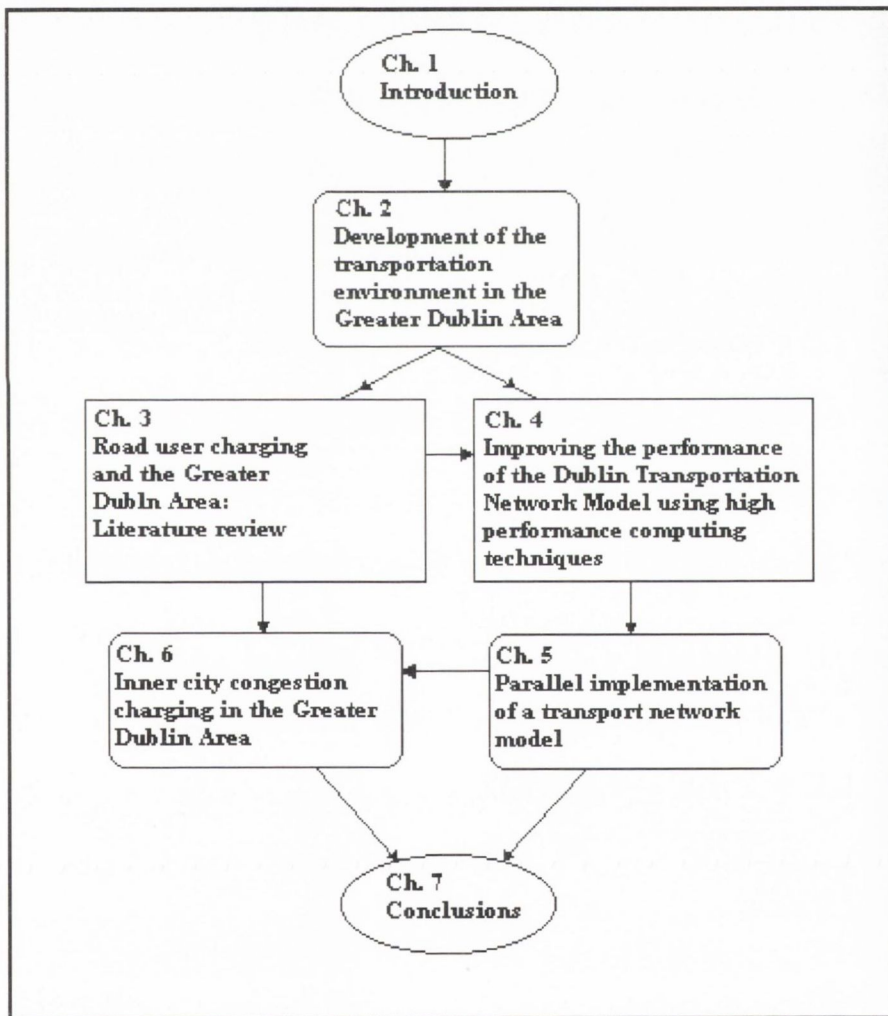
#### **1.4 Outline of Thesis**

There are 7 chapters in this thesis, which relate to each other in the manner shown in Figure 1.1. In **Chapter 2** a detailed analysis of the development of the transportation environment of the GDA is presented. This includes a look at the planning processes that have shaped the GDA and a brief discussion of the proposed new planning and administrative structures. Details of Dublin's development over the last 15 years is presented and some conclusions drawn as why the changes have taken place in the manner that they have. Future development of the GDA is then discussed with particular reference to traffic congestion and the existing transportation deficit.

In **Chapter 3** the development of road user charging is described through a review of international experience and literature. Basic design criteria and the practical aspects of design and implementation are discussed briefly. The political and social aspects of road user charging and the effectiveness of road user charging schemes are then discussed in the context of pre and post implementation case reviews and studies. The price elasticity of user demand is then examined in the context of this international experience and literature. Finally, relevant conclusions are drawn from this review in relation to the application of a road user charging scheme in the GDA.

**Chapter 4** has two distinct parts. In the first part the Dublin Transportation Network Model (DTNM) is described. The computationally intensive assignment and simulation processes involved in the DTNM are analysed further and an investigation into the results of this analysis are discussed. This leads to a review of the assignment and, in particular, the shortest paths problem with respect to the DTNM and the assignment.

In the second part of Chapter 4 a literature review of parallel programming techniques, methodologies and considerations is undertaken. This review also describes parallel hardware and software architectures and examines performance analysis techniques used for parallel systems. Finally, an analysis of parallel processing in traffic assignment and shortest path theory is presented with a view to reprogramming the assignment and shortest path sections of the DTNM.



**Figure 1.1: Thesis outline.**

In *Chapter 5* the parallel implementation of the sequential transport network model is presented. The methodology is presented outlining: a brief background to the application of parallel processing techniques to the transport model; the parallel

platform used to run the new parallel model; the parallelisation strategy for the sequential model based on the comprehensive review in *Chapter 4*; and, the programming implementation of the single-source one-to-all shortest path algorithm in parallel. The results from this parallel implementation are then presented and discussed with respect to the convergence, accuracy and performance of the new parallel model when compared to that of the sequential model.

*Chapter 6* presents a series of tests using the new parallel model, which look at congestion charging in the GDA with particular reference to an inner city congestion charging scheme. The chapter is laid out in much the same way as *Chapter 5* with a methodology detailing: the cordon model inputs; the charge levels and elasticity estimates used based on the comprehensive literature review in *Chapter 3*; the demand function and generalised cost equations; and, specific details pertaining to the highway network and cordon application. The results and discussion of this testing is then presented using a number of parameters including distance travelled, average speed, queue lengths and travel time.

*Chapter 7* contains final discussion and conclusions of the work and some suggestions for further work.

## *Chapter 2*

# *Development of the Transportation Environment in the Greater Dublin Area*

## 2.1 Introduction

The Greater Dublin Area, or GDA, is defined as the city of Dublin and the areas defined by the county councils of Dun Laoghaire-Rathdown, Fingal and South Dublin. The catchment area of these county councils includes parts of county Dublin and parts of counties Meath, Wicklow and Kildare, as seen in Figure 2.1 below. A map outlining the boundary of the GDA can also be seen in Figure 2.2 in the context of the Metropolitan Area. This Section will examine the development of the transportation environment in Ireland, with specific reference to the GDA over the past four decades, leading to the transportation deficit that exists in the GDA at present. To do this,



Figure 2.1: Basic map of the Greater Dublin Area showing the Metropolitan Area, the simulation area boundary and the inner city cordon.

strategic proposals and reports that have been produced influencing the development of transportation in the GDA will be assessed with a view to examining how successfully they have been implemented and how this implementation has affected the development of the transportation environment in the GDA. Future proposals influencing the development of transportation will also be described and evaluated with reference to the situation at present.

There are a number of reports that have been produced describing plans and strategies for the future development of the Dublin area (as defined in previous studies to the Dublin Transportation Initiative) and the GDA. These have tried to map out the development of the GDA through transportation planning and land use management. As a result, these reports have had a major influence over the development of transportation in the GDA and the resulting transportation deficit that exists at present. However, attempts have been made to rectify this deficit, including new legislation, a transportation strategy for implementation between 2000 and 2016 and the production of strategic planning guidelines, which outline proper planning and sustainable procedures for both transportation and land use development.

## ***2.2 Transport planning in the Greater Dublin Area***

### *2.2.1 Introduction*

Ireland's planning system has undergone major changes in recent times. However, it is the old system that has influenced development in Ireland and the Greater Dublin Area (GDA) over the past four decades. The old planning system, 1963–1999, will be discussed with reference to the GDA as background material for other areas of the thesis.

The planning of transportation and land use has been going through a period of change since the introduction of the new Planning Bill in 1999, which outlines sustainable procedures for proper planning practice. The Bill and its implications will be discussed in the context of the changing transportation planning infrastructure of the GDA.



While the new Planning Bill represents the catalyst and outline for change, the publication of the Strategic Planning Guidelines (SPG) (Brady Shipman Martin et al., 1999), see Figure 2.2, has provided a co-ordinated settlement plan for the GDA and the

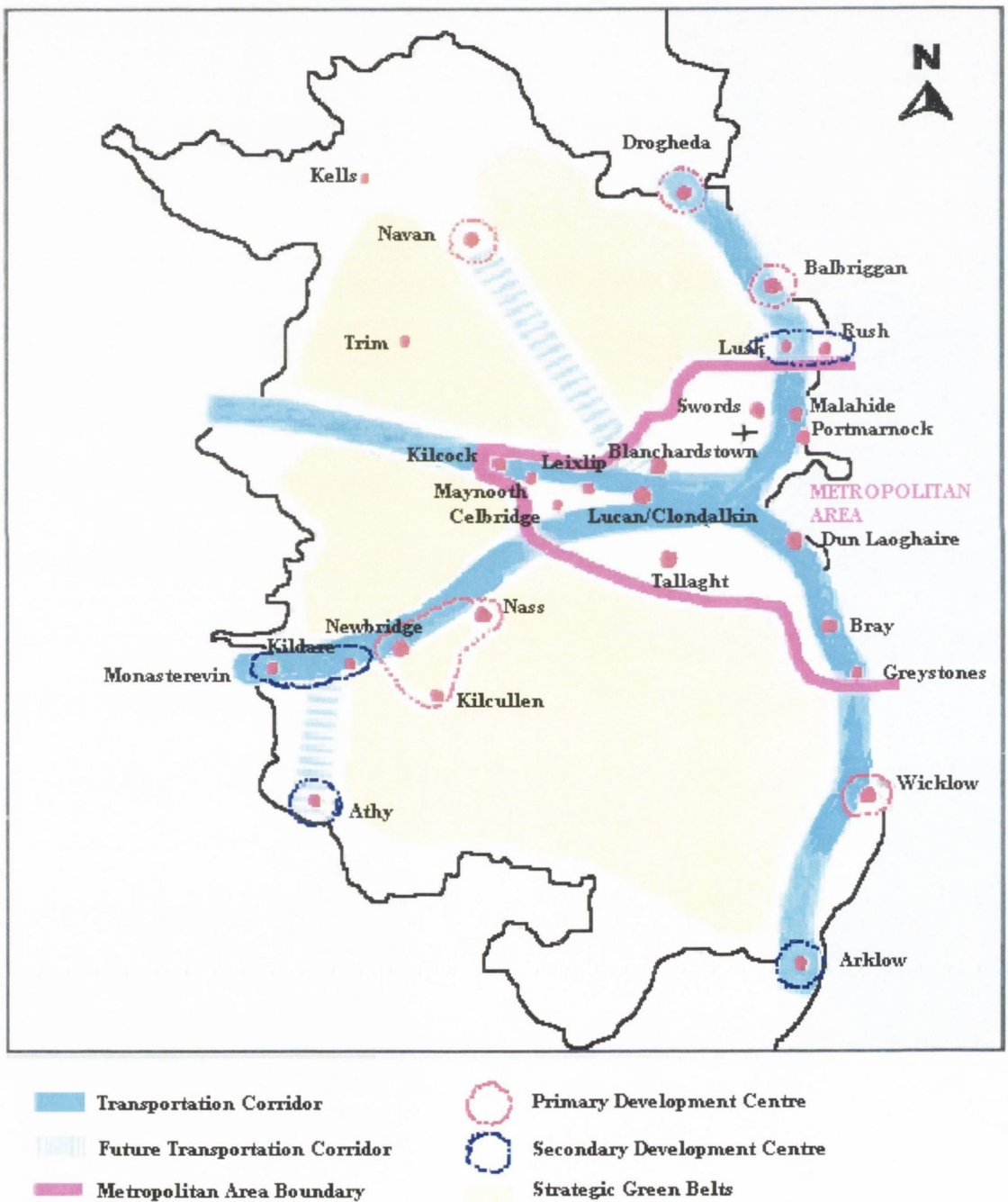


Figure 2.2: Strategic map of the Greater Dublin Area (GDA) taken from the Strategic Planning Guidelines (Brady Shipman Martin et al., 1999)

Platform for Change: Strategy 2000 – 2016’ (DTO, 2000) will be discussed in later Sections of this chapter.

Finally, the establishment of a new authority with regulatory control over strategic land use and transport issues in the GDA and the implications of deregulation will be discussed.

### 2.2.2 *The Irish planning system, 1963 – 1999*

The planning system in Ireland has been predominantly based on the Local Government (Planning and Development) Act of 1963. This legislation came into effect on the 1<sup>st</sup> of October 1964 and defined for the first time a “*statutory development control system*” (DoE, 7/12/2001). As such, the basis of Irish urban planning lies in land use zoning and development control (MacLaran, 1993). The Minister for the Environment and Local Government is responsible for all planning legislation and regulations. Many have been introduced since the 1963 Act and predominantly reflect the changing economical environment, the introduction of the statutory and independent planning appeals system (An Bord Pleanála) and legislation associated with our membership of the European Union. The various acts resulting from these changes are collectively known as the *Planning and Development Acts*, while the regulations are known as the *Planning and Development Regulations*.

The planning system in Ireland is run by 88 local planning authorities whose principal responsibilities include:

- Producing development plans;
- Assessing planning applications and accepting or refusing permission;
- Appeals against planning decisions; and,
- Planning enforcement and exempt development.

The GDA is predominantly made up of four local planning authorities, three of which are county councils with the other being a city council. They are Dublin City Council,

Fingal County Council, South Dublin County Council and DunLaoghaire-Rathdown County Council.

Regulation of development is controlled by the Development Plan. Each local planning authority must produce a development plan every five years, though this time schedule isn't always adhered to because of delays associated with public participation. The development plan lays out policies and controls for transportation and land use development in the local authority area. The authority must consider the development plan and its objectives when making planning permission decisions. The public is notified if a draft development plan is being produced or if an existing plan is being revised. The draft is open to public criticism for at least three months and the public is entitled to make submissions to the authority on it. These submissions must be considered before the development plan is adopted. The draft must also be sent to various statutory and voluntary institutions for specialist advice as well as to the adjoining local authorities.

The decision to grant or refuse planning permission lies with the planning authority. The authority must assess applications for planning permission by considering them in the context of the development plan, i.e., the authorities' objectives for land use, road improvements, development and renewal of obsolete areas, and, amenities. Any observations or submissions made by other bodies (e.g., the Environmental Protection Agency) are also taken into account.

Certain developments, as defined in the First Schedule of the European Communities (Environmental Impact Assessment) Regulations, 1989, must be examined for the likely environmental impacts before planning permission can be granted. Applications for these types of development must submit an Environmental Impact Statement (EIS) when submitting for planning permission. In the case of a refusal for planning permission the authority is required to inform the applicant of the reasons for the refusal.

In the case of a refusal an appeal can be made to An Bord Pleanála (the Planning Appeals Board). Usually a planning authority will decide on a planning application

within two months unless more information is requested from the applicant to clarify the planning application. The decision to make an appeal must be made within a month of the refusal. The judgement made by An Bord Pleanala may only be appealed in the High Court on a point of law within two months of An Bord Pleanala's decision to refuse permission. The High Court will not reassess the planning merits of the case.

The planning system described here, based largely on the 1963 Act, initially emphasised the active involvement of planning authorities in development (MacLaran, 1993). This did not happen because of a lack of public finance. In practice, the planning system depends on the private sector to put forward potential development projects. As such, development plans must set out a format for growth and prosperity that encourages the private sector to participate in the development of transport and land use.

While the Irish Planning System appears to be a very democratic one, allowing the right of appeal to third parties even though they may not have a direct association with the development proposal, in practice the process of public participation tends towards a public relations exercise (MacLaran, 1993) that is cumbersome and extremely frustrating for potential developers. Unfortunately, development plans also take longer than the recommended five years to review. Dublin's present development plan was introduced in 1999 but the previous Plan was adopted in 1991, eight years earlier.

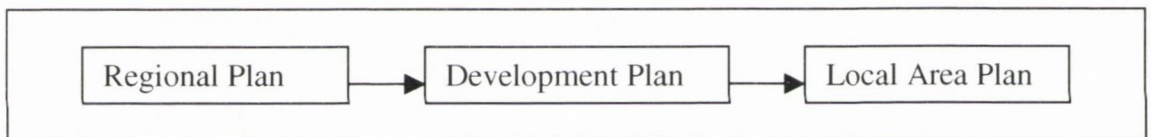
Due to the influence of the private sector over development, the Irish Planning System can be considered flexible to pressure for or against development. As such it lacks long term strategic values. The continental approach to planning enables more long-term strategic thinking with a more proactive, rather than reactive, approach.

### 2.2.3 *The Irish planning system, 1999 –*

Nine Planning and Development Acts have been enacted between 1963 and 1999. The new Planning Bill has revised and consolidated these Acts while also incorporating a sustainable development ethos into the Irish planning system. The production of development plans will remain at the core of sustainable and proper planning but with some changes:

- Planning authorities must now produce a Development Plan every six years;
- The plan shall be consistent with any national plans or strategies relating to proper planning and sustainable development and shall set out “*an overall strategy for the proper planning and sustainable development of the area*” (DEHLG, 2000); and,
- The plan must address not just land use scenarios but also the provision of infrastructure including transport, energy and communications facilities.

The development plan is subservient to the regional plan while setting out the background for any local area plans, see Figure 2.3. Each of the plans has been given statutory recognition providing a consistent planning process.



**Figure 2.3: The set-up for a more cohesive planning system as set out in the Local Government (Planning and Development) Act, 2000 (DEHLG, 2000).**

Local area plans have been given an increased status to allow a more consistent relationship between development plans and individual development proposals. Local plans can be produced for any area and when such an area lies between two different planning authorities then these authorities can co-operate amongst each other. An Environmental Impact Statement (EIS) must also be produced outlining any likely effects on the environment through the implementation of a local area plan.

Regional plans must incorporate a number of issues. However, only some of the more important characteristics pertaining to this work are mentioned here:

- The SPG (Brady Shipman Martin et al., 1999) for the GDA prepared for Dublin City Council and the County Councils of Dun Laoghaire-Rathdown, Fingal, Kildare, Meath, South Dublin and Wicklow in conjunction with Dublin Regional Authority and the Mid-East Regional Authority have effect as if

made under the Local Government (Planning and Development) Act, 2000 (DEHLG, 2000);

- They are to be developed for a period of time between 12 and 20 years;
- They should provide a long-term strategic planning framework for the development of the region in accordance with the principles of proper planning and sustainable development; and,
- Regional plans must also address the following matters, amongst others:
  - Transportation, including public transportation;
  - The location of industrial and commercial development;
  - Economic and employment trends; and,
  - Projected population trends and settlement and housing strategies.

These regional guidelines will allow for a better cohesion of development plans. The Bill (DEHLG, 2000) sets out the purposes for which objectives may be indicated in a development plan. *'Infrastructure and Transport'* represent a separate section showing the controls that a development plan can have over the provision of transportation. The inclusion of transport planning in the Bill provides for the co-ordinated planning of transport and land use by ***“promoting sustainable settlement and transportation strategies in urban and rural areas”*** (DEHLG, 2000). Some of the more important controls include:

- Reserving land for transport networks, including roads, rail, etc.;
- Facilitating the provision of sustainable integrated transport, public transport and road traffic systems and promoting the development of local transport plans;

- Improving convenience and safety for users of the transport network, including pedestrians and cyclists;
- Establishment of public rights of way and extinguishment of public and private rights of way; and,
- Providing for the management and control of traffic, including the provision and control of parking areas.

With respect to the GDA, local authorities will be obliged to consider the regional plan produced by Dublin Regional Authority when reviewing development plans. This should produce a cohesive planning system allowing consistent planning decisions resulting in proper planning and sustainable development.

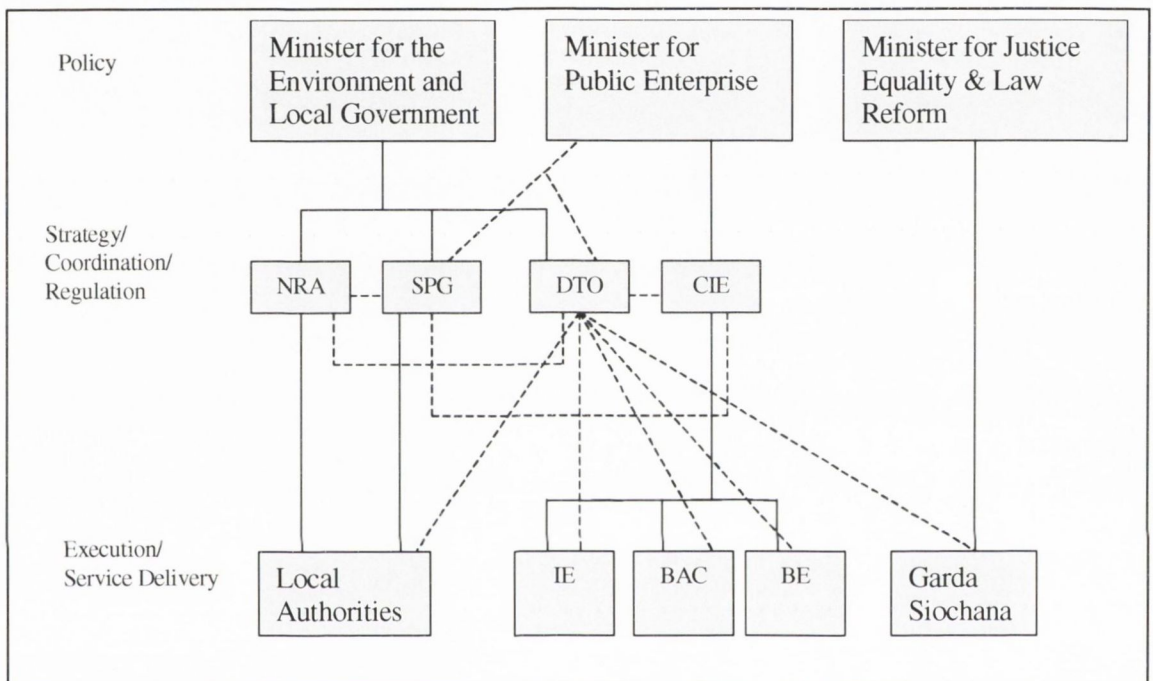
In addition to regional planning, a National Spatial Strategy (DEHLG, 2002) has been prepared “*to shape balanced regional development in Ireland for the next 20 years*” (Minister Martin Cullen TD, 2002). The National Development Plan 2000–2006 is consistent with this long-term strategy. Also, any regional or development plans must work within the framework of the National Spatial Strategy.

#### 2.2.4 Existing institutional arrangements for transport and land use in the GDA

At present, there are a number of different bodies in operation that have responsibilities in relation to land-use and transportation in the Greater Dublin Area (GDA). These can be seen in Figure 2.4, which shows how the differing bodies interact with each other.

The responsibility for transportation in Ireland is divided between two government departments: the Department of the Environment and Local Government and the Department of Transport. The Department of Public Enterprise held responsibilities until the Department of Transport was established in 2002. The Department of the Environment and Local Government has responsibility for policy, legislation and financing related to general traffic and road infrastructure. The strategy and co-ordination to undertake this policy is controlled by the National Roads Authority (NRA), the SPG Steering Committee and the DTO.

The NRA was established under the Roads Act of 1993 and became active on the 1<sup>st</sup> of January 1994. The NRA has a planning, co-ordination and funding role in relation to national roads and has the responsibility of providing a safe and efficient network operating to international standards. However, implementation is primarily exercised



**Figure 2.4: Existing transport planning institutional arrangements for the Greater Dublin Area (DELG & DPE, 2001).**

through city and county local authorities. While the NRA oversees any work associated with the national road network, it is the local authorities that have responsibility for the maintenance and improvement of local roads. These works must also be financed with the authorities' own resources although State grants may be provided in certain circumstances.

The DTO is a corporate body established on the 9<sup>th</sup> of November 1995, under the Local Government Services (Corporate Services) Act, 1971. Apart from the control over certain traffic management related funding, the DTO must pursue its mandate through voluntary arrangements and consensus. The Department of the Environment and Local Government sponsor the DTO and it has a Steering Committee comprised of a Local



Authority Committee and Consultative Panel. These are made up of a number of elected councillors, representatives from key transport agencies, representatives of both the Department of the Environment and Local Government and the Department of Public Enterprise, and, other outside expertise. The responsibilities of the DTO are to:

- Co-ordinate and monitor the implementation of the DTI Strategy (see the DTI Final Report (Steer Davis Gleave, 1994a));
- Review the activities of agencies involved in project implementation to ensure that their actions are consistent with the DTI Strategy; and to,
- Review and update the Strategy every five years.

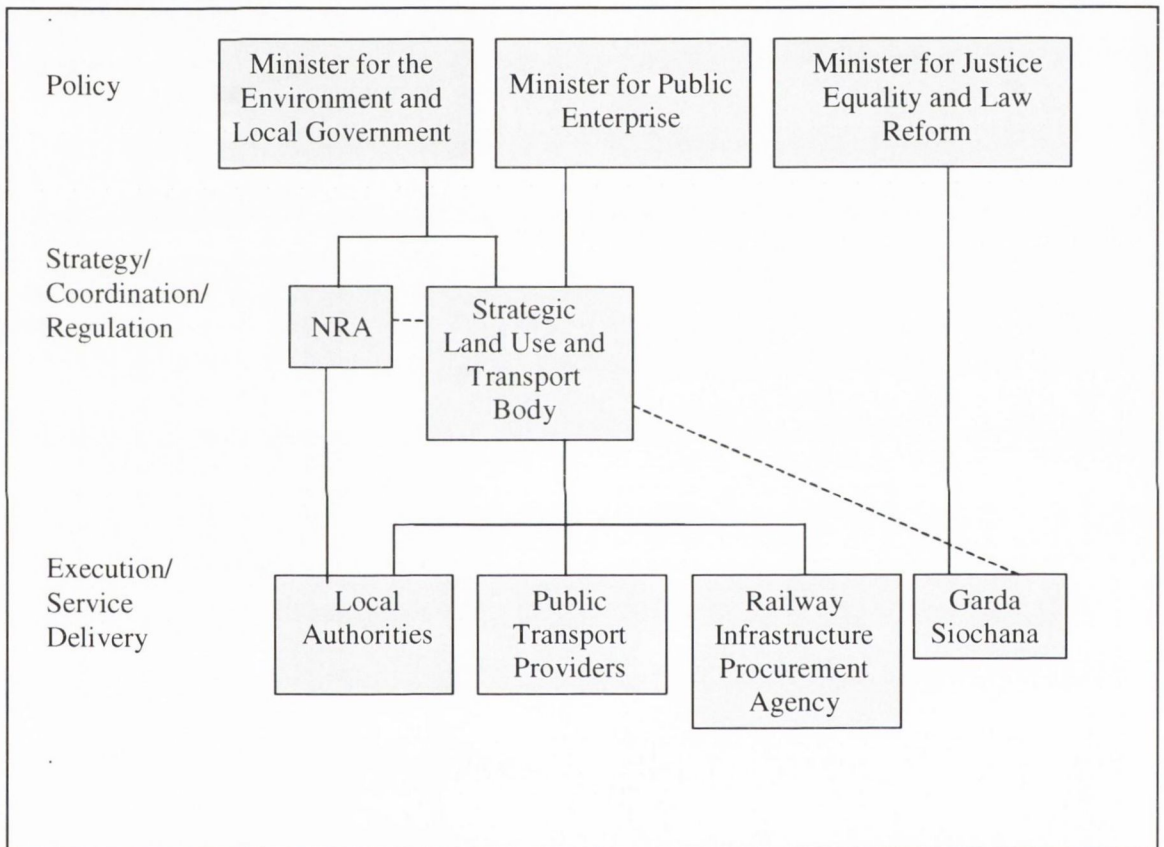
The DTO is managed by a Director who exercises his duties with the help of a team of specialists in the areas of transportation planning, transportation modelling, traffic management and land-use planning.

Separate arrangements exist to monitor the implementation of the SPG (Brady Shipman Martin et al., 1999) involving a Steering Group with representatives from the local authorities, both Departments, Coras Iompair Eireann (CIE) and the DTO, a technical group, and, a full-time technical director.

The Department of Public Enterprise has responsibility for policy, legislation and financing related to public transport. It is a shareholder in the State transport company CIE. CIE was established under the Transport Act of 1950. Under the Transport Act, 1986, CIE was required to form a holding company and three major operating subsidiaries. CIE and three operating subsidiaries: Iarnrod Eireann (Irish Rail); Bus Atha Cliath (Dublin Bus) and Bus Eireann (Irish Bus), are responsible for the provision of the vast majority of public transport services in Ireland. Bus Atha Cliath provides passenger services for the city and county of Dublin. Iarnrod Eireann provides mainline passenger services, the Dublin Area Rapid Transit (DART) service and suburban services in the GDA. Bus Eireann provides a bus passenger service throughout the country.

### 2.2.5 Proposed institutional arrangements for transport and land use in the GDA

The Department of the Environment and Local Government and the Department of Transport have outlined 'New Institutional Arrangements for Land Use and Transport in the Greater Dublin Area' in a consultation paper published in March 2001. Figure 2.5, when compared to Figure 2.4, shows the proposed changes to the institutional



**Figure 2.5: Proposed transport planning institutional arrangements for the Greater Dublin Area (DoELG & DPE, 2001).**

infrastructure in the GDA. It can be seen from Figure 2.5 that two main changes are proposed:

- The establishment of a strategic land use and transport body, or GDA Authority; and,
- Public transport institutional and regulatory changes, i.e., deregulation of the public transport sector to facilitate a transport market.

These changes have been proposed because it has been found that the existing institutional arrangements are not capable of managing the development of the infrastructure and public transport services proposed in the National Development Plan 2000 – 2006 and the DTO’s long term transportation strategy, A Platform for Change (DoELG & DPE, 2001). The Public Transport Partnership Forum (October 2000), the DTO (2000) and a range of other bodies have all agreed that the existing institutional arrangements are not adequate and that a new single statutory body with the following powers is required:

- The body should have responsibility for land use and transport planning in the GDA;
- The body should deal with transport and land use strategic planning and co-ordination in an integrated way;
- The body should have adequate financial powers and resources; and,
- The body should be democratically accountable at a political level.

The Department of Public Enterprise published a consultation paper on deregulation of the public transport sector in August 2000. In this paper the main impetus for the change in policy, i.e., deregulation, is described as the *“urgent need to deal with traffic congestion problems in Dublin”* (DPE, 2000). Public transport is also recognised as an important tool in promoting sustainable development and addressing social inclusion. Figure 2.5 shows the new institutional changes necessary to facilitate the development of a market in public transport.

The Strategic Land Use and Transport Body seen in Figure 2.5 will represent an independent public transport regulatory body that will:

- Regulate the bus market through franchising and licensing;
- Negotiate public service contracts and award public transport service franchises; and,

- Be responsible for the regulation of matters arising from public private partnerships and EU legislation.

These institutional and regulatory measures are being put in place to help facilitate the strategic and policy driven objectives for the GDA. The National Development Plan 2000–2006 sets out these general objectives as being:

- To address the projected growth in traffic through transport infrastructure and facilities investment and demand management measures;
- To make commuter travel by private car more unattractive thereby reducing congestion and emissions;
- To increase accessibility and support sustainable development; and,
- To reflect commuter travel patterns by providing for a spatial distribution of public transport in line with the SPG.

### ***2.3 Transportation development in the GDA prior to the Dublin Transportation Initiative (DTI)***

#### ***2.3.1 Introduction***

Dublin's inner city streets developed over a long period of time. The City is over 1000 years old and was never bombed during the Second World War, unlike some of its European counterparts. As such, it retains its narrow streets. By 1914 Dublin City had a network of just over 100 kilometres of electrically powered trams. After the First World War this network was replaced by the modern motorised bus, which was deemed to be more flexible and forward looking at the time (Killen, 1992).

Many reports have been published outlining planning proposals for the Dublin area. However, these proposals have rarely ever been implemented in full. Some of the reports concentrate on land-use while others look more closely at transport issues.

Considering the interconnections and dependency of each discipline on the other the important sections of the reports will be summarised in chronological order.

### 2.3.2 *The Myles Wright Report (1967)*

The first attempts at planning the development of the Dublin area were produced in the 1960s. However, these reports did not take into account the fact that the areas of transportation and land-use are closely related. For example, the Schaechterle traffic plan was produced in 1965. It endorsed a comprehensive road widening scheme in the inner city, which was proposed by Dublin Corporation to help keep traffic moving. Two years later the Myles Wright Report was produced. This proposed a population shift away from the inner city to relieve excess commuter traffic. As a result of the contradiction between the transport oriented Schaechterle traffic plan and the land-use oriented Myles Wright Report the majority of the Schaechterle plan was not implemented and Dublin's inner city streets remain narrow to this day. However, many buildings became derelict along streets earmarked for widening. When this widening never came to fruition the Government then had to offer the private sector costly incentives to rejuvenate the derelict sites (MacLaran, 1993). A number of improvements were made though, including the East Link Bridge and the East – West routes along the Royal and Grand Canals.

The Advisory Regional Plan for the Dublin Region, or the Myles Wright Report (1967), led to the development of four new towns to the West of Central Dublin. These new towns were based on small existing settlements. The four new towns proposed were:

- Tallaght – Saggart;
- Clondalkin – Milltown;
- Palmerstown – Lucan; and,
- Blanchardstown – Mulhuddart.

A high population growth was predicted and it was planned that these towns would cater for the development necessary as a result of this growth. They were planned as self-sufficient settlements providing employment, services, housing and amenities to their inhabitants. These proposals were not implemented as planned. Even so, the Myles Wright Report (1967) has had a large influence over the current shape of the Metropolitan Region of the GDA, see Figure 2.1 and Figure 2.2.

The Advisory Plan proposed an expansion of the City limits 10 miles to the West over a period of 20 years, taking account of a predicted population growth of 275,000 over the same period. Wright (1967) states that “*new building developments around Dublin should be at low density*”. This low density development was considered a cheaper alternative to high density construction where land use was plentiful. In summary, the main strategic points of the Report were that:

- The majority of the new population should live in the new self-sufficient settlements;
- These towns and the city centre should be far enough apart so as to spread the traffic between them; and,
- Trips between the new communities would take place over routes not suffering from congestion.

These proposals were to encourage shorter journeys by keeping employment closer to home and, in so doing, reduce traffic congestion and the growing pressure on Central Dublin. Wright recommended heavy use of the bus for commuting between the new towns and between the new towns and Central Dublin because of the low density housing proposed, and encouraged the development of infrastructure for the bus and private car as a means of dealing with any unknown needs in the future (Wright, 1967).

As MacLaran surmises (1993), due to the proximity of the new settlements to the city and the manner in which the development was carried out the likelihood of their achieving self-sufficiency was negligible. Planning was based on accommodating the car, which only encouraged urban sprawl. No organisation was given the role of

overseeing the development of these new towns. Instead, implementation of the proposals was based on private sector investment (MacLaran, 1993). Consequently, the proposals have not come to fruition, as is demonstrated by the fact that none of the new towns have developed town centres and only Tallaght received its shopping centre in 1990. Also, the proposed green belts to be observed between the new settlements have not been enforced and a continuous suburban area now envelops what should have been independent self-sufficient towns.

### 2.3.3 *Subsequent studies*

The Dublin Transportation Study was produced in 1971 because it was generally accepted that transport plans in the Myles Wright Report lacked detail (Steer Davis Gleave et al., 1992). Though little of this report was realised proposals included an outer motorway system and scaled down proposals for road development in the inner city. Upgrading of both rail and bus commuter services were also endorsed.

The Dublin Rail Rapid Transit Study (DRRTS) was produced by Vorhees and Associates in 1975. This CIE commissioned study proposed a rail system serving 47 stations, mostly underground, with a city centre terminus near O'Connell Bridge and an extensive feeder bus system linked to the train system. In 1978 a Transport Consultative Commission was set up to examine the DRRTS system.

The resulting report (Transport Consultative Commission, 1980) found that the DRRTS proposal was too expensive and the only part that was implemented was the electrification of the Howth/Bray Dublin Area Rapid Transit (DART) line, which was opened in 1984. The DART is considered to be one of the most successful pieces of public transport development in recent times. Unfortunately perhaps, instead of advocating the implementation of all of the DRRTS proposals, the Transport Consultative Commission recommended improved transport links for Tallaght in the form of a bus service in the short term and with complementary features such as bus lanes and parking restrictions to follow.

The Eastern Regional Development Organisation (ERDO), set up in 1969, published its first report in 1985. The organisation, made up of representatives from local authorities including Dublin County Borough, County Dublin, Dun Laoghaire and counties Kildare Meath and Wicklow, examined population growth and its effects in the Eastern Region up to 2011. To do this, it looked at land-use options, paying little attention to transportation effects. The report focused on regional effects and this brought it into conflict with those trying to reinforce inner city development (Dublin Crisis Conference Committee, 1986). The body appears to have been totally isolated from transport planning but did acknowledge the importance of locating additional population near transport facilities.

The next phase of planning has developed out of the Dublin Transport Initiative (DTI). Much of the transport policy arising from the DTI is still currently in use and its importance warrants a separate subsection.

## **2.4 The Dublin Transportation Initiative (DTI)**

### *2.4.1 Introduction*

The DTI began in 1988 when the Minister for the Environment commissioned a review of transport planning. The first phase of the DTI declared that:

*“It has become clear that what is now required is more than a once off study, rather a transportation planning process which will permit the regular review of the transportation strategy and investment or implementation programme and allow on-going assessments of the impact of land-use and other policies on transportation and vice versa”* (Steer Davis Gleave, 1992).

As such, the DTI heralded a new approach to planning in the Dublin region. The study area, or DTI Study Area, is shown in Figure 2.2. It covered the City and County of Dublin and parts of Counties Kildare, Meath and Wicklow including all of the built-up areas making up Dublin City (Steer Davis Gleave, 1994a).



The first phase of the DTI included an analysis of previous reports on land-use and transportation planning. These included the Myles Wright Report (Wright, 1967), the Dublin Transportation Study (1971), the Central Dublin Traffic Plan (Morgan et al., 1973), the Dublin Rail Rapid Transit Study (Vorhees et al., 1975) and the Transport Consultative Commission (1980). The study also incorporated all the relevant groups including CIE, the relevant Government departments and local authorities and advocated strong public involvement. This was the first time that such an open forum had been advocated and then implemented, which led to real enthusiasm that progress would be made in the right direction. Phase one of the DTI proceeded to set out an approach for the second phase, saying that it should:

- Be all embracing of transport modes;
- Include assessment of the interaction between transportation and other policies;
- Provide a process;
- Provide a long term strategy (2011) and a five year investment and implementation programme with short term policies for the period between 1994 and 1999 called the Operational Programme for Transport;
- Look at the question of enforcement;
- Consider funding in detail; and,
- Give special consideration to the peripheral suburbs and new towns, the Inner City and the access needs of both the Port and Airport for both freight and public passengers.

In 1992 phase two of the DTI began under the aegis of the objectives laid out in phase one above. The most important objectives for phase two of the Initiative were set out to be:

- The drafting of a long and short term transportation strategy; and,

- The organisation of a continuous planning process.

After two years of intensive negotiation and consultation between all interested parties and the general public, the DTI Final Report, laying out Dublin transportation plans, was published in May 1994 (Steer Davis Gleave, 1994a).

#### 2.4.2 DTI methodology

The first step in producing a DTI Strategy was to draft a Vision Statement and to use this to set out the guiding principles for the rest of the Study. The Vision Statement provided the core for more detailed objectives which the Strategy would be designed to fulfil. The Vision Statement can be seen in Figure 2.6.

- **Dublin as a leading European City:** Competing and co-operating successfully, civilised, literate and vibrant.
- **The National Capital:** Proud of its history and heritage, its unique character conserved, a fit setting for Government and national centres of excellence.
- **A Metropolitan Region** with a strong, growing and diverse city, town and rural economy, based on a skilled and adaptable work force.
- **A Living City – Region** on a human scale, accessible to all: at its heart, a city to serve its people and communities and to meet their aspirations for an improving quality of life.

**Figure 2.6: The Vision for Dublin, which guided the Dublin Transportation Initiative, had four facets (Steer Davies Gleave, 1994a).**

From this Vision Statement potential transport strategies were developed. Agencies within the Study Area then put forward possible schemes and plans, which were examined in the context of policy and organised into eight Themes. These can be seen in Figure 2.7.

The eight themes were not examined with a view to finding a winning theme; rather it was used to establish the contributions of individual policy options towards meeting the objectives extrapolated from the Vision Statement. Once the strengths and weaknesses of the different policy options were assessed a fundamental strategy would then be adopted using different elements of the policies that performed well.

1. **Do Minimum** – The benchmark situation for 2011 for comparison purposes, comprising only those schemes that have already been given the go-ahead.
2. **Restraint** – To reduce car usage through road pricing and other transport management measures.
3. **Making Better Use of Existing Assets** – Low cost approach to increase public transport and road usage through traffic control measures and new rail stations.
4. **Incremental Road Development** – Investment in road infrastructure through the completion of major road schemes and some minor developments.
5. **Extended Road Development** – A more expansive package of road investment measures including all those from Theme 4 plus other extensive road developments.
6. **Environmental Package** – Focusing on pedestrianisation, priority for cyclists and related measures such as traffic calming.
7. **Enhanced Public Transport Investment** – Including QBCs, Light Rail Transit (LRT) and improvement to heavy rail suburban services.
8. **Extended Public Transport Investment** – Provide a city-wide heavy rail network plus extensions to existing lines and a new city centre public transport interchange.

**Figure 2.7: The DTI theme based approach (Steer Davies Gleave, 1994a).**

The main criteria were:

- (1) Employment and the Regional Economy;

- (2) Quality of Life;
- (3) International and National Context;
- (4) Development of the City and the Region; and,
- (5) Efficiency in Implementation.

A full list of the objectives/criteria and their sub-criteria can be seen in Appendix 1. The themes were then tested using a traffic model and their performances were compared to the “*Do Minimum*” theme and to each other.

Due to the interaction of land use and transportation it was also decided that a number of land use scenarios would be assessed so as to give further insight into the policy options. However, the DTI did not have any land use planning powers and no long term land use plans existed for them to base their transportation strategy upon. As a result, the DTI realised it would have to produce a number of possible land use development scenarios for the future. These scenarios can be seen in Figure 2.8.

**Scenario A: Urban Concentration** – Household and employment growth concentrated in inner city and surrounding towns with control over development in surrounding areas.

**Scenario B: Growth Poles** – Household growth in western towns with employment growth in the city centre, the western towns plus Swords, Maynooth and the Bray area.

**Scenario C: Dispersal** – Household growth along radial routes and in North County Dublin. Employment growth along the C-Ring motorway and outside of urban areas

**Figure 2.8: The DTI land use scenarios (Steer Davies Gleave, 1994a).**

The DTI also stressed that these scenarios were developed as hypothesis of how the GDA would develop from a land use perspective over the projected period and that they were not meant to replace any existing short or medium term land use strategies.

These three scenarios were chosen through a process of socio-economic and demographic analysis (Steer Davies Gleave, 1994b). Once chosen the scenarios were developed for low, medium and high growth potential between 1991 and 2011, taking population and household data into account. These can be seen in Table 2.1.

**Table 2.1: Low, medium and high growth between 1991 and 2011 for the DTI Study Area (Steer Davies Gleave, 1994b).**

	<i>Low</i>	<i>Medium</i>	<i>High</i>
<b>Population (millions)</b>	0.986	1.203	1.221
<b>Household size (persons)</b>	2.96	2.86	2.31
<b>New households</b>	77,000	77,000	77,000

In order to calibrate the three scenarios the ‘central figure’ of 77,000 households was allocated over the Study Area. In actuality, for an average size of 2.84 persons, approximately 80,000 new households would have been accommodated, while 63,000 new households would have been accommodated for an average household size of 2.84 persons, by the year 2011 (Steer Davies Gleave, 1994b).

The eight themes were assessed using a method of multi-criteria analysis. The criteria used in the analysis are those mentioned previously and listed in Appendix 1. Each theme was analysed by assessing and then awarding its sub-criteria a proportionally positive or negative score depending on their performance in comparison to the other themes. As such, the criteria score differently depending on what theme is being assessed. For example, Table 2.2 shows a number of sub-criteria, which are of some interest, and how they performed under the different themes; the letter ‘N’ represents ‘Not Given’.

With reference to the criteria mentioned in Table 2.2, Theme 2, ‘Restraint’, performs very well. However, Themes 3 and 4 also have slightly positive effects and Themes 7 and 8 perform very well in certain aspects as well. This led the DTI to believe that a multi-faceted approach to development in the Study Area was necessary, taking the

positive aspects from each of the themes to produce an effective and comprehensive strategy. Theme 5, though performing well on a number of criteria, was rejected because of its conflict with the higher level strategic aims of the DTI to reduce car dependence and reinforce the city centre.

While a network of quality bus corridors (QBCs) was assumed necessary as the basis for a strong public transport foundation, further analysis was undertaken to assess the potential benefits of light and heavy rail. A cost benefit analysis (CBA) was conducted to compare the economic performance of the following two options:

- A heavy rail system serving Dublin Airport and suburbs to the North, South and East of the City Centre; and,
- A light rail network serving Tallaght, Dundrum and Ballymun to the West, South and North of the City Centre, respectively.

**Table 2.2: Performance of some sub-criteria of interest over the eight themes numbered above (Steer Davies Gleave, 1994a).**

Sub-criteria	1	2	3	4	5	6	7	8
<b>2.1 Improved travel reliability</b>	---	++	+	+	++	N	+++	+++
<b>2.4 Reduce environmental impacts of transport</b>	--	++	+	+	++	N	++	++
<b>3.3 Improve access to and from the GDA</b>	---	+	+	+	+++	N	+	++
<b>5.1 Contain financial needs</b>	+++	+++	+++	++	---	+++	--	--
<b>5.6 Financial health of transport operating costs</b>	--	+++	++	N	N	+	++	N

The cost benefit ratio of each option was produced using the bus network as a baseline comparison. The results produced showed that heavy rail, at 0.4 : 1, was the least

favourable option when compared to the proposed light rail project, which produced a ratio of 1.76 : 1. This led to the inclusion of an LRT in the DTI final strategy.

#### 2.4.3 *Recommended strategy of the DTI*

The DTI Final Report (Steer Davies Gleave, 1994a) set out a vision of Dublin. It saw Dublin as a leading European Capital, economically competitive, historically proud, accessible to all, a growing and diverse city based on a skilled and adaptable work force. In order to achieve these aims the DTI laid out a number of objectives in order to encourage economic regeneration and development throughout the Study Area and to help maintain and reinforce the city centre as the country's prime commercial and cultural centre. The strategy contains a number of measures to be implemented up to the year 2011 and can be summarised as being to:

- Improve public transport and not to let the private car dominate the transport system;
- Emphasise the movement of people and goods, not just vehicles, and in the case of the ports and airport, address the access requirements for both freight and public passengers;
- Bring greater equity to the transport system, improving accessibility for all and taking account of the real needs of disadvantaged; and,
- Provide an integrated approach to transportation and land-use.

The DTI Strategy also considers a number of traffic management measures including traffic calming in residential areas and environmental traffic management in the city, especially in the city centre. There are strong implications for how traffic is managed in the city centre as a result of the implementation of the DTI Strategy. Phase two of the DTI states that;

*“Nothing less than its (city centre) conversion from a car and heavy vehicle dominated, pedestrian unfriendly location to a vibrant, efficient area is being proposed”* (Steer Davies Gleave, 1994a).

The Initiative advocates a better, more efficient use of resources such as public transport, the SCATS traffic-light control system and land-use planning so that major infrastructural changes can be implemented in a more receptive environment:

- Public transport should be up-graded and given more priority through the introduction of 10 Quality Bus Corridors (QBCs);
- SCATS, which covers the city centre area roughly between the Royal and Grand canals, should be extended to cover a greater area of the city; and,
- High-density land development should be encouraged along priority public transport routes and in areas earmarked for future development.

In the long term the Initiative proposed new and improved public transport links, especially between the new towns of Tallaght, Lucan–Palmerstown, Clondalkin and Blanchardstown to the west and between these towns and the city centre. Other proposals include:

- A light rail network to be developed and implemented along routes of high trip demand such as those from Tallaght to the city centre and from Sandyford to St. Stephen’s Green, with a potential link to the airport also being considered important;
- The DART being extended and the DART and suburban rail services being up-graded; and,
- Road improvement projects such as the M50 C–ring, the Port Access Tunnel and the South Eastern Motorway.

The Port Access Tunnel will allow HGVs to enter and exit the port area with comparative ease while cars are tolled for the privilege of getting direct access to the



city centre via the tunnel. A number of other traffic management measures are also proposed; these include:

- Park and ride facilities to be set up at 9 locations near the C-ring where good public transport facilities are present to transport people from outside the C-ring into the city centre;
- Parking restrictions in the city centre to be enforced more stringently with the addition of clamping and additional resources to spend on towing and clamping equipment; and,
- Long term parking to be deterred by restructuring parking charges and reducing the number of non-residential off-street parking facilities.

Environmental traffic management measures including the development of environmental traffic cells, pedestrianised zones and cycle paths are included in the proposal to encourage a more equitable sharing of road space by all modes of transport so as to make the roads and streets of Dublin more safe and pleasant for people and to reduce the environmental impact of traffic (Steer Davies Gleave, 1994a).

The DTI Strategy recommended a co-ordinated land-use policy for the Dublin metropolitan region and local authorities were encouraged to take account of the DTI recommendations when next reviewing their development plans. Unfortunately, the DTI Strategy was based on future land-use scenarios that had no influence over the way in which the Dublin metropolitan region would really develop. Due to the close interaction and dependency of land-use and transportation development, the DTI Strategy could never provide a definite long-term solution unless a land-use strategy was designed to complement it. Only in recent times has this been recognised in Ireland with the publication of the SPG (Brady Shipman Martin et al., 1999).

Finally, the DTI Strategy went on to call for new institutional arrangements for the running of the transportation planning process and to oversee implementation of the strategy itself. This led to the establishment of the DTO, as a corporate body, in November 1995 to co-ordinate and monitor the implementation of the DTI Strategy.

#### 2.4.4 *Transportation modelling and the DTI*

This Section provides a summary of the transportation network modelling process used in the DTI for the Study Area and for Dublin's recent transportation plans. This modelling process is known as the Dublin Transportation Network Model (DTNM). An understanding of this modelling process provides a basic foundation from which this thesis will try to develop further ideas and from which it will also attempt to improve certain aspects. Areas within the process will also be used in conjunction with the new developments to assess transport management scenarios in the GDA.

Sections of the DTNM are discussed in more technical detail in Chapter 4, Section 4.2.

##### **2.4.4.1 The DTI model approach**

The DTI used a conventional approach to transportation modelling. This approach can be broken down into four stages:

- Trip generation and trip attraction;
- Trip distribution;
- Modal split; and,
- Trip Assignment.

Each stage has its own sub-models and the four stages are generally carried out in the order shown above. This approach can be applied at a strategic level for medium to long term planning or it can be used to support detailed scheme design and implementation issues. The DTI used two software suites, SATURN (Simulation and Assignment of Transport in Urban Road Networks) and SATCHMO (SATURN Transport Choice Model).

For the purposes of modelling the Study Area was sub-divided into a series of zones. There were two levels of zoning; the coarser, or strategic, level with 58 zones used to assess land use, car ownership and trip growth changes and the finer, or tactical, level

with 367 zones used to analyse mode choice, assignment problems and the actual traffic flow and distribution within the Study Area. These finer zones were developed by subdividing the coarser zones and are made up of wards, District Electoral Divisions (DEDs) and six external peripheral counties representing the rest of the Ireland.

#### **2.4.4.2 Data collection**

The strategic zoning level is used in conjunction with the land use scenarios mentioned earlier. These scenarios are used to provide the basic information from which the trip generation data is derived, and hence, the trip generation model must also be applied at this coarser level. The model can be described as being local in that the strategic zones are addressed individually with growth factors being determined from the land use scenarios and demographic data. The model also contains disaggregate household data particularly pertinent to car ownership and car availability.

This data was obtained from census material, which was then aggregated to the strategic 58 zone level as it was originally produced at ward level. Data produced in this way included information about population, employment, average occupancy of households, cars per 1000 population and cars per household. The census data also allows strategic zonal totals of journeys to work and education on foot and by bicycle. However, the DTI assumed that these ‘slow modes’ share the same journey characteristics as those travelling by bus. In effect, ‘slow mode’ transport characteristics were not catered for in modal choice. Therefore, the resulting modelling process did not allow for analysis of potential pedestrian or cyclist policies.

Two base year origin – destination matrices were produced using the peak (08:00 – 09:00) and off peak periods (14:00 – 15:00) for light goods vehicles and cars and for Heavy Goods Vehicles (HGVs) to represent actual trip patterns. To do this, three sets of origin – destination surveys were undertaken:

- Road side interviews;
- On-board bus interviews; and,

- Rail passenger surveys.

During the road side interviews movements were recorded for cars and light goods vehicles and for HGVs. A traffic count was also taken at the same time as the interview allowing the origin – destination information to be scaled up (grossed) to take account of travel during the entire model period and not just during the interview period. Once this was done at each site a number of logic checks were undertaken to ensure the accuracy of the travel patterns recorded. These logic checks involved assessing a number of routes in more detail and comparing the predicted route totals with the actual totals. Once these proved to be comparable an assumption was made that similar scaling factors would apply to all other routes. Then the information from each site was pooled together to form the seed matrices. Any unobserved data was then in-filled on these matrices using a software package.

Bus and rail surveys were also taken. The bus questionnaire was designed to be compatible with both the roadside interviews and a separate network planning review being undertaken by Bus Atha Cliath (BAC). The questionnaire was handed to passengers boarding each bus for self-completion and then collected as passengers egressed at their specific stops. The boarding stops were recorded by noting the code number of the questionnaire and the alighting stops recorded by storing the collected questionnaires in a file specific to that stop.

The total number of passengers for each route was found by scaling up the results from completed questionnaires using the total number of people travelling on each route extracted from Electronic Ticketing Machines (ETMs) supplied by BAC. In this way a trip matrix was produced for the period that the matrix was to represent.

The rail survey was performed in a similar manner to the bus survey with forms handed out at suburban rail stations and the finished survey forms being collected at destination stations. The survey was held very close to Iarnrod Eireann's annual boarding/alighting counts. These were used to scale up the survey data received to produce a rail matrix. A matrix was also produced for non-car trip makers depending on the availability of a car, or not, for both bus and rail passengers.

### 2.4.4.3 DTI trip generation and future year matrices

This section of the model is used to forecast the increase in trips originating from a zone (generation) and destined to a zone (attraction) between the base year and the forecast years (Steer Davies Gleave, 1994b). Growth factors were developed for the 'car available' and the 'car non-available' matrices. The two models were then applied over the peak and off peak periods. Trip forecasts were then produced for each zone under each of the three land use scenarios described.

The trip generation model is applied at the 58 strategic zone level. The peak trip generation model predominantly uses the growth in work trips and school trips. The working population is calculated by taking the total zonal unemployment from the 15 – 65 age group. The off peak generation and attraction models are similar, with the former using origin data and the latter, destination data; for more detail see the DTI Technical Report Volume 2 (Steer Davies Gleave, 1994b). Land use is treated as an input to the attraction model in that there is a separate attraction model for separate travel purposes. Growth in trips is based on the growth in employment within each strategic zone. Trip totals calculated from trip generation models are taken to be more accurate than trip attraction totals because trip generation models use growth factors calculated using more reliable household data. Due to this, trip attraction totals are adjusted to fall into line with the equivalent trip generation totals.

Future year trip matrices are produced using growth factors and the application of the Furness Iteration Technique. Separate growth factors were produced for the number of trips originating and terminating in each zone. The Furness method is a doubly constrained technique that allows the estimate of a trip matrix that satisfies both growth factors by balancing them against each other.

The output future year trip matrix between zone 'i' and 'j',  $T_{ij}$ , is given by:

$$T_{ij} = t_{ij} A_i B_j \quad \text{Equation 2.1}$$

where:

- $t_{ij}$  is the base year trip matrix (input matrix); and,
- $A_i, B_j$  are the balancing factors incorporating the growth factors for each zone.

For the first iteration the balancing factors are derived by dividing the future trip end total by the current trip and total, i.e:

$$A_i = \frac{O_i}{o_i} \quad \text{Equation 2.2}$$

$$B_j = \frac{D_j}{d_j} \quad \text{Equation 2.3}$$

where  $O_i, D_j$  are zonal origin and destination totals for the future year and  $o_i, d_j$  are similar for the base.

The method works by alternatively applying each balancing factor. For the first two iterations this is done by setting  $B_j$  equal to zero and solving for  $A_i$  and vice versa. For further iterations the balancing factor is recalculated from the results of the previous iteration, iteration number two in this case, with  $O_i$  as previously defined and  $o_i$  being the origins resultant from the previous iteration. After a small number of iterations this process converges to some predefined measure, usually within 5% of target values.

#### 2.4.4.4 Generalised cost and mode split

Generalised cost is of key importance in modelling mode choice in the DTI. It is assumed that journeys are made for the utility (value) taken from activities that are performed at their destination. However, when making such a journey the traveller incurs the disutility of expense, time spent travelling, etc. Generalised cost represents this disutility. The zone to zone generalised cost used by the DTI is given by the two formula:

$$GC_{\text{car}} = a_1IVT + a_2WK + a_3EX + a_4VOC + a_5PK + a_6 \quad \text{Equation 2.4}$$

$$GC_{pt} = a_1^1 IVT + a_2^1 WK + a_3^1 WT + a_4^1 FARE + a_6^1 \quad \text{Equation 2.5}$$

where:

- IVT is the in-vehicle travel time;
- WK is the walk time;
- WT is the waiting time for public transport;
- EX is the excess time while finding parking;
- VOC are the vehicle operating costs;
- FARE are the public transport fares;
- PK refers to the parking charges for the private car; and,
- $a_1$  to  $a_6$  are parameters used as both conversion factors and weights to express generalised costs in common units.

The generalised cost can be expressed in monetary or time units. If the generalised cost is measured in monetary terms ' $a_1$ ' can be interpreted as the perceived value of IVT. Usually ' $a_2$ ' and ' $a_3$ ' are approximately twice that of ' $a_1$ ', making the perceived value of walking and waiting times twice that of IVT. Parameters ' $a_4$ ' and ' $a_5$ ' convert values of fare into their equivalent values in time while ' $a_6$ ' is the mode specific constant representing the inherent costs of a mode.

The DTI have used a hierarchical or nested logic form of mode choice model. The first, higher nest, choice that is made is between the private car and public transport using the logit formula below:

$$P_{car} = \frac{e^{-\beta GC_{car}}}{e^{-\beta GC_{car}} + e^{-\beta GC_{pt}}}$$

$$\text{Equation 2.6}$$

$$P_{pt} = 1 - P_{car}$$

where:

- GC is the generalised cost; and,
- $\beta$  is a scaling parameter representing the effects of mode choices to changes in the GC by each mode.

This finds the number of travellers with a car available that will use either their car or public transport. The same type of formula is then applied to public transport to see what the split is between bus and rail for both 'car available' and 'car non available' public transport users, i.e., two  $\beta$  scaling parameters are needed.

$$P_{bus} = \frac{e^{-\beta GC_{bus}}}{e^{-\beta GC_{bus}} + e^{-\beta GC_{rail}}}$$

Equation 2.7

$$P_{rail} = 1 - P_{bus}$$

For the higher nest choice it is necessary to form an equation representing the cost of public transportation as a combination of both bus and rail costs, which are each found using the public transport generalised cost equation given previously. To be consistent with the logit models the logsum average of the respective modes costs was used:

$$GC_{pt} = -\frac{1}{\beta} \ln(e^{-\beta GC_{rail}} + e^{-\beta GC_{bus}})$$

Equation 2.8

#### 2.4.4.5 Assignment

The assignment area of the DTNM will be looked at in greater detail at a later stage because it is within this area that this thesis attempts to provide performance improvement through the use of an IBM SP2 supercomputer that is jointly owned by Trinity College Dublin and Queens University Belfast.

The DTI had two macro models available to them: CONTRAM (**C**ONTinuous **T**raffic **A**ssignment **M**odel) and SATURN, and chose to use SATURN as its platform to represent the detailed interaction between junction delay and assignment. SATURN



uses two models to give realistic assignment values. The simulation model uses cyclic flow profiles to represent groups of vehicles moving across the network. In this way the delays in one approach to a junction, which depend on one or more of the other approaches plus itself, can be represented more accurately. However, it needs information about the flow on each link to estimate capacity, queues and delays. To get this information an assignment model is used, which loads a trip matrix onto the network and obtains an estimate of these flows. Using SATURN this can be done with either Wardrop's equilibrium assignment or stochastic user equilibrium assignment.

Passenger assignment is also performed for the public transport networks and this is carried out by SATCHMO, a complimentary package to SATURN. SATCHMO considers a number of attributes associated with public transport route choice including: cost, in-vehicle time, transfer time, walking and waiting time. Passengers then choose the best route in terms of generalised costs. When the assignment of the private car and public transport have been completed there may be some difference between the travel times and costs calculated and those used in the mode split models described earlier. Therefore, an equilibration process is used to balance the estimated generalised costs and those found through assignment. When equilibrium conditions are reached then no user can reduce his or her generalised costs by changing either their route or their mode.

#### *2.4.5 Implementation of the DTI measures*

In the year following the publication of the DTI strategy its implementation was already falling behind schedule. However, the DTO, which had been planned for in the DTI, emerged in 1995 to strategically monitor the transportation planning process. Many of the transportation plans were to be completed by 1999, as laid out in the Operational Programme for Transport (1994–1999). However, due, in part, to the division of responsibility for the implementation of the DTI strategy amongst a number of agencies, the DTO were unable to avoid substantial delays. Table 2.3, below, shows estimated finish dates for some of the DTI schemes.

The implementation of traffic policy, including traffic management and the implementation of QBCs and parking policy, is the responsibility of the Office of the

Director of Traffic, established in 1997 in Dublin Corporation (now Dublin City Council). Traffic cells have been introduced in and around the city centre. A number of turns have also been banned except for public transport, bicycles, motorbikes and taxis. These measures have been put in place to reroute through traffic away from the city centre and to give public transport and pedestrians greater priority in the area.

**Table 2.3: Differences in estimated finishing dates for some of the DTI strategy measures from 1994 to 1998.**

<i>Scheme</i>	<i>Estimate of finish date, 1994 (Steer Davies Gleave, 1994)</i>	<i>Estimate of finish date, 1998 (DTO, 1998)</i>	<i>Estimate of finish date, 2003 (DTO, 2001)</i>
Southern Cross Route (M50)	1999	2001	2003
South – Eastern Motorway (M50)	1999	2003	2006
Northern Port Access Tunnel	1999	2003	2006
Tallaght/Connolly (LRT)	1999	2002	2004
Sandyford/St. Stephens Green (LRT)	1999	2003	2005
St. Stephens Green/Ballymun (LRT)	2001	Post 2003	Cancelled

Only four of the initially proposed ten QBCs were completed by the proposed finish date in 1999. However, five new QBCs opened in early 2001 and the DTO immediately reported notable increases in passenger volumes at that time. By the end of November 2001 nine of the DTI's eleven QBCs had been opened (DTO, 2001). Bus journey times have also been brought down along QBC routes. This can be seen along the Lucan QBC where buses travel 20% more quickly and along the Rathfarnham QBC where there is a 55% improvement, according to figures from Dublin Bus (QBC Steering Committee, 2002). The QBCs that have been established for a longer period of time, i.e., Malahide, Lucan and the N11, were also reported to be carrying more than

twice as many people as a lane of general traffic at the design speed of 22kph (DTO, 2001). A further 15 QBC schemes have been planned for completion over the next number of years; see the QBC Report to the DTO QBC Steering Committee (QBC Steering Committee, 2002).

Parking policy has encouraged short-term parking. The city has been divided into five zones with charges being proportional to their proximity to the city centre and the areas of highest demand. The current highest rate in 2003 is Euro 2.90 per hour in the city centre. Enforcement of parking policy has been vastly improved with wheel clamping being introduced in 1998 and an increased level of towing of illegally parked vehicles throughout the city region. Consequently, revenues are up and there are fewer illegally parked vehicles clogging up the streets.

The Light Rail Project Office (LRPO), now the Railway Procurement Agency (RPA) following the enactment of the Transport (Railway Infrastructure) Bill in 2001, is responsible for the implementation of the LUAS light rail system project. The RPA is an independent statutory State body whose main responsibilities will be the procurement of the Metro and light rail projects through Public Private Partnership (PPP). The LUAS scheme is considerably behind schedule because of two main factors: the expansive public protest at the required infrastructural changes and the resulting public consultation process; and, the prolonged bureaucracy and long periods of indecision resulting from a change in national government during the planning stages. These factors plus a number of engineering logistic difficulties have meant that the planned finishing date for the lines from Tallaght to Connolly Station and from Sandyford Industrial Estate to St Stephen's Green have now been pushed back to 2004 and 2005, respectively.

The NRA is responsible, through the respective councils, for major road schemes. All of the proposed major road schemes, including the Dublin Port Access Tunnel, Southern Cross Route and South-Eastern Motorway have all been substantially delayed by the planning appeals process. In late 2000, the Southern Cross Route was under construction while the Port Tunnel and the South-Eastern Motorway were at the tender

stage (DTO, 2000). However, over 125km of bicycle track and 2300 new bicycle parking spaces have been made available and 2000 extra spaces have been provided at suburban rail stations (DTO, 2000).

## 2.5 *Dublin's development since the DTI began*

### 2.5.1 *Economic growth and population change*

There has been a huge surge in traffic growth in the Dublin metropolitan region since the DTI began. This traffic growth has resulted from the economic boom in Ireland, and in the Dublin region in particular, since the DTI Final Report was published in 1994. This has, in part, attributed to delays in implementing sections of the DTI strategy.

**Table 2.4: Factors affecting transport demand in the DTI Area (DTO, 1998).**

DTI Area	1991 Base	2001 Forecast	1996 Base	1999 Forecast
Population (m.)	1.14	1.18	1.18	1.19
Households ('000)	343	378	379	393
Unemployment rate	15.8%	17.0%	12.4%	11.6%
Car ownership per 1000	248	288	300	350
GNP Growth	Base	49%	42%	70%

Table 2.4 and Table 2.5 show a number of factors that affect transport demand over the DTI Study Area and over the GDA, as defined by the DTO. It can be seen that for the DTI Study Area 49% GNP growth was forecast by 2001 yet by 1996 42% had already been achieved with an updated forecast of 70% having been predicted for the year 1999. The actual GDP growth in 1999 for the GDA was 79%, exceeding the DTI forecast for that year, and growth of 260% was also forecast for the year 2016.

The SPG (Brady Shipman Martin et al., 1999) also indicate that the 2016 forecast of population, see Table 2.5, is conservative by at least 100,000 and that the figure for the number of households is also conservative by as much as 10%, see Section 2.6.2. This growth has outstripped all expectations, which is simply illustrated by the following points:

- The predicted population for 2001 was exceeded in 1997;
- GDP grew by 79% as opposed to a predicted 38% between 1991 and 1999;
- The DTI forecast a 17% rate of unemployment in 2001 and 12% in 2011 while the rate in 2000 was a little under 6%;
- Car ownership per 1000 population was 342 in 1999, fast approaching the EU average of 450 cars per 1000 population, compared to a predicted forecast of 288 for the year 2001.

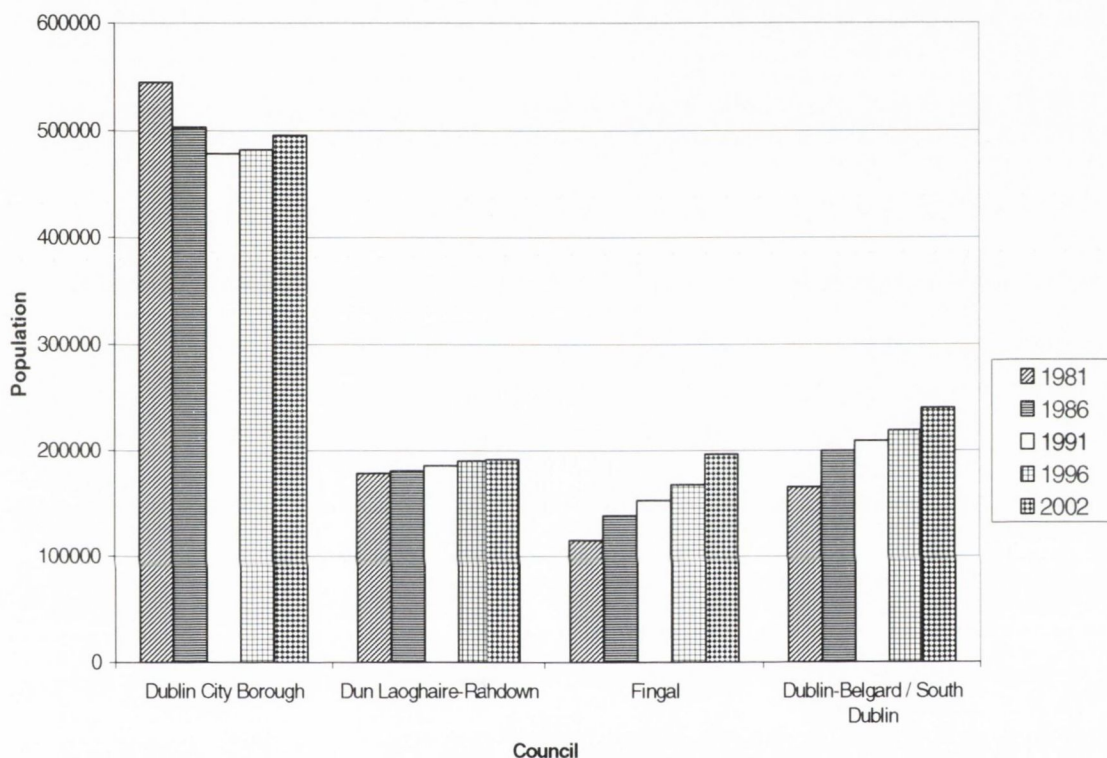
**Table 2.5: Factors affecting transport demand in the GDA (DTO, 2000).**

Greater Dublin Area	1991 Base	1996 Base	1999 Actual	2016 Forecast
Population (m.)	1.35	1.41	1.46	1.75
Households ('000)	402	446	521	675
Unemployment rate	16%	12%	6%	5%
Car ownership per 1000	247	292	342	480
GDP Growth since 1991	Base	42%	79%	260%

This growth can be attributed to a couple of important factors. Firstly, government policy was aimed towards inward investment and attracting foreign investment through tax incentives and other enticements. Secondly, the European Union gave large amounts of structural funding allowing extensive infrastructural development. This has facilitated investment from foreign sources and encouraged economic growth in the

region. In 1996 direct overseas investment, co-ordinated through the Irish Development Authority (IDA), made up 45% of all manufacturing jobs and 71% of exports.

The economic growth has in turn led to a large influx of people living and working in



**Figure 2.9: Changes in the population of council areas in the Dublin metropolitan region between 1981 and 2002 (Central Statistics Office, 2002).**

the GDA. There has also been a large increase in the construction of office space to accommodate the market place (Heaney, 2001). A large quantity of this office space, approximately 40% between 1996 and 2000, is located in proximity to the M50 ring road (HOK 1997 – 2000). As well as this, a number of shopping complexes such as The Square, Liffey Valley and Blanchardstown have been constructed on major junctions of the M50 corridor. Due to the peripheral location of these facilities and their close proximity to the M50 the major means of access is with the private car. This, and the attraction of these facilities, has resulted in an increase in the number of people travelling in the GDA, especially during peak periods.

Population location has also effected trip demand in the GDA. Figure 2.9 shows an increase in population in the peripheral areas of Dun Laoghaire–Rathdown, Fingal and Dublin-Belgard / South Dublin while an overall decline is found up until the 1996 census in Dublin City Borough. However, as a direct result of the implementation of schemes relating to the DTI objectives for the inner city, to create a vibrant sustainable inner city environment, this population trend has changed and the population is now increasing in the Dublin City Borough as well as in peripheral regions.

Overall, the general population has increased and there has been a population shift from the inner city to more peripheral suburban areas, particularly in the Fingal and Dublin-Belgard / South Dublin council areas. Dublin city has a dispersed central core and low intensity of use with a policy of limiting height for inner city developments (DEGW, 2000). This, along with rising land prices in the inner city, has encouraged development to take place in suburban areas. This is demonstrated by the increase in population in the Fingal and Dublin-Belgard / South Dublin Council areas where there is ample room for low-density development. This low-density development has in turn been made more favourable through rising car ownership, leading to increases in the number of trips being made and to the distance travelled during peak periods.

### 2.5.2 *Increased travel demand*

The DTI's predictions for growth in the demand for travel have been dramatically exceeded. This is not surprising considering the unprecedented growth in the Irish economy, growth that the DTI could not have foreseen. Table 2.6 shows the peak hour and off-peak hour total trip demand figures for 1991 and 2001 from the initial 1991 DTI model. It also shows the equivalent figures for 1996, 1997 and 1999 from the 1996 model updated by the DTO. These figures show that the demand for travel in 1997 exceeded the predicted levels for 2001 by 16%.

Table 2.7 shows the equivalent figures taken from the latest DTO revision of the DTI model, published by the DTO (2000), and gives a forecast for peak and off-peak demand in 2016. The 1997 peak and off-peak travel demand figures of 234,000 and 130,000 have been revised upwards to 250,000 and 157,000 respectively. Total peak

hour trips have increased by 78,000, or 45%, between 1991 and 1997, with the private car being accountable for approximately 71,000 of these. This represents an increase in the use of the private car from 64-72% over the same time period. The increase in congestion is demonstrated by the fact that the average journey time by car has increased from 31 minutes in 1991 to 43 minutes in 1997 (DTO, 2000).

By 2016 total peak hour travel demand is predicted to increase by a further 95% over 1997 levels, up to 488,000 trips. In 1999 peak hour demand had already increased by more than 13% over the 1997 level, exceeding the necessary growth rate to achieve the 2016 forecast. Total off-peak demand is forecast to be approximately 256,000 in 2016. This total is higher than peak hour demand in 1997.

**Table 2.6: Transport trip demand and growth from the DTI and the DTO (DTO, 1998).**

	Peak Hour Demand		Off-Peak Demand	
	<i>Person trips</i>	<i>Growth since 1991</i>	<i>Person trips</i>	<i>Growth since 1991</i>
1991 Base	172,000		107,000	
2001 Forecast	206,000	20%	126,000	18%
1996 Base	227,000	32%	125,000	17%
1997 Actual	234,000	36%	130,000	22%
1999 Forecast	240,000	40%	134,000	25%

The changing environment in the GDA, resulting from unprecedented economic growth, higher demand for travel and changes in demographic trends, has led to changes in patterns of travel. A comparison has been made by the DTO (1998) of journey destinations in the morning peak period for surveys carried out in 1991 and 1997. In 1991 the city centre was the primary destination. However, the 1997 survey showed that:



- While the city centre remains the main destination other destinations such as Clondalkin/Tallaght have had a fourfold increase in trips made by car in the interim and the area is the fastest growing destination in the GDA;
- Dublin Airport is a major destination for trips and there is an axis of travel demand between the western towns of Blanchardstown, Clondalkin, Tallaght and Dublin Airport; and,
- The other areas of increasing significance include Sandyford in the South and the northern fringes of Dublin City.

These trends demonstrate the spread of development investment in low-density suburban areas and of the substantial increase in population in the western towns of Tallaght, Lucan, Clondalkin and Blanchardstown.

**Table 2.7: Transport trip demand and growth from the DTO (DTO, 2000).**

	Peak Hour Demand		Off-Peak Demand	
	<i>Person trips</i>	<i>Growth since 1991</i>	<i>Person trips</i>	<i>Growth since 1991</i>
1991 Base	172,000		107,000	
1997 Actual	250,000	45%	157,000	47%
1999 Actual	283,000	65%	179,000	67%
2016 Forecast	488,000	184%	256,000	139%

In accordance with other trends of growth in the GDA, private car ownership levels have increased as the economy has prospered. The DTI used a base value of 248 cars per 1000 of population in 1991 for the DTI study area while a value of 288 cars per 1000 of population was forecast for 2001, see Table 2.4. However, when an updated model was produced, using 1996 as the base year, it was found that the value of 288 cars had already been surpassed with a value of 300 cars per 1000 of population being used as the new base. The new forecast for the DTI study area was given as 350 cars

per 1000 of population to be achieved by 1999. As is seen in Table 2.5, the actual value achieved in 1999 across the GDA, a larger area defined by the DTO, was 342 cars per 1000 of population with a predicted value of 480 cars per 1000 of population by the year 2016.

**Table 2.8: Private cars licensed for the first time (Central Statistics Office (CSO), 2002).**

New Private Cars					
<i>2001</i>	160,908	<i>Jan-Jul 2002</i>	118,938	<i>Jul 2002</i>	14,386
<i>2000</i>	225,269	<i>Jan-Jul 2001</i>	129,641	<i>Jul 2001</i>	14,558
<i>1999</i>	170,322	<i>Jan-Jul 2000</i>	178,225	<i>Jul 2000</i>	19,728

The value of 480 cars per 1000 of population is well beyond the EU average of 450 cars per 1000 of population. However, this forecast was produced by the DTO during an economic boom that has since come to an end. From Table 2.8 it is apparent that the number of new private cars licensed for the first time in 2001 was 160,908, a decrease of 29% on the figure for 2000 but only a decrease of 6% when compared to 1999.

There were 118,938 new private cars licensed in the first seven months of 2002, a decrease of 8% on the corresponding period of the previous year. In July 2002 and 2001, 14,386 and 14,558 new private cars were licensed compared to a figure of 19,728 for July 2000, approximately a 27% reduction. These figures have reflected recent trends in the economy. As such, it is the state of the economy that will determine how quickly Ireland will progress towards the EU average for car ownership.

### 2.5.3 *Dublin's transportation deficit*

A key objective of the DTI strategy (Steer Davies Gleave, 1994a) was to achieve an effective modal shift away from the use of the private car in favour of public transport. This would improve the transportation environment in the GDA by reducing

congestion, travel times, pollution, and improving accessibility, especially in the inner city region.

The DTO defines a transportation deficit as ‘any increase in person-trips by private car over that forecast by the Dublin Transportation Initiative (Steer Davies Gleave, 1994a)’ (DTO, 1998). Simplistically, this is viewed as a combination of an increase in the demand for travel and a reduction in the market share of public transportation. As the previous Sections have made clear, the DTI strategy fell badly short of initial infrastructural and strategic targets, see Section 2.4.5. The level of demand for travel has also increased by more than double what was predicted for 2001 in the 1991 DTI model forecast, and this was achieved in 1999, two years prior to the forecast year.

In addition to these factors, institutional and regulatory inefficiencies, discussed in Section 2.2, and a cumbersome and time consuming planning appeals system have resulted in numerous delays to infrastructural projects and unrealistic strategic planning targets for land use and transportation development. These trends have resulted in a substantial transportation deficit in the GDA.

## ***2.6 Future development in the Greater Dublin Area***

### *2.6.1 The Strategic Planning Guidelines for the Greater Dublin Area*

The SPG Report (Brady Shipman Martin et al., 1999) was produced as a “***framework for integrated land use and transportation for the sustainable development of the GDA up to the year 2011***” (SPGO (Strategic Planning Guidelines Office), 2002). The SPG were prepared for the councils of Dublin City, Dun Laoghaire-Rathdown, Fingal, Kildare, Meath, South Dublin and Wicklow and for the Dept. of the Environment and Local Government, in conjunction with Dublin Regional Authority and the Mid-East Regional Authority and produced by a number of consultants including: Brady Shipman Martin, Kirk McClure Morton, Fitzpatrick Associates and Colin Buchanan and Partners. This expertise was later formalised into the SPGO whose mandate has been to monitor and update the SPG. The key elements of the SPG include:

- The encouragement of transport modes other than the private car, including a better public transport system within the Metropolitan Area, see Figure 2.2, and good links between Development Centres and the Metropolitan Area to reduce car congestion and to better integrate these centres into the economy of the GDA;
- Incorporation of the principles of sustainable development as established in the National Sustainable Development Strategy (DEHLG, 2002);
- Extensive strategic green belt areas where development will be encouraged to meet local needs and not those arising from commuting; and,
- More concentrated development within the Metropolitan Area and within Development Centres along transportation corridors in the Hinterland, see Figure 2.2.

The SPG considered a number of relevant points from the National Sustainable Development Strategy 1997–2002 (DEHLG, 2002) including:

- The minimisation of potential growth in transport demand as a leading consideration in land use planning;
- Attempts to provide more sustainable and environmentally acceptable alternatives to private car transport, especially public non-motorised transport;
- The promotion of closer co-ordination between transport and land use planning; and,
- An intensification of the implementation of the DTI and making the public more aware of the negative effects of vehicle transport.

The SPG attempted to make economic growth predictions in 1999. However, as a result of the events of 2001 the Economic and Social Research Institute (ESRI) made alternative predictions in their Medium Term Review 2001–2007 that the SPG have

now adopted, see Table 2.9. Both forecasts predict an economic decline with growth moderating at approximately 4.5% (SPGO, 2002).

The 2011 forecast for the population of the GDA in the SPG Report (Brady Shipman Martin et al., 1999) was 1.65 million, see Table 2.10. However, this was increased to 1.76 million in the SPG Review and Update (SPGO, 2002) because of long term Central Statistics Office (CSO) predictions. It was changed once again by the CSO 'M1F2' high or low net national in-migration forecast to between 1.8 and 1.73 million. During the present economic uncertainty migration rates may fall however, in which case the original 2011 population forecast in the SPG (Brady Shipman Martin et al., 1999) may prove accurate.

**Table 2.9: ESRI alternative growth predictions (ESRI, 2001).**

<b>Benchmark Forecast</b>	<b>2000</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>
GNP	9.9	6.0	5.4	4.4	4.0	4.0	4.4	4.3
Employment Growth	4.8	4.3	2.2	1.2	1.0	1.4	1.7	1.8
Unemployment Rate	4.3	3.8	3.6	4.3	5.3	5.8	5.7	5.5
<b>Slowdown Forecast</b>	<b>2000</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>
GNP	9.9	6.0	1.8	4.2	5.1	5.3	6.1	4.6
Employment Growth	4.8	4.3	-0.9	0.2	2.0	2.9	3.4	2.4
Unemployment Rate	4.3	3.8	6.3	7.6	7.4	6.5	5.1	4.6

The household figures in Table 2.10 are based on a prediction that the average persons per household will continue to drop to 2.5 by 2011. This trend along with increasing population gives rise to a high demand for housing.

The SPG are used as the basis for planning in the GDA by all Government Departments and Agencies, including Local Authorities, when preparing development plans. The priority of the SPG is to control where this new development will take place in the

GDA, see Figure 2.2 for an outline of the proposed ‘Development Centres’. This will allow for a better balance between public and private transport by securing growth in a limited number of locations along existing or future transportation corridors and consolidating growth in the Metropolitan area so as to allow for future population expansion (Brady Shipman Martin et al., 1999).

**Table 2.10: Population and household predictions for the Greater Dublin Area (Brady, Shipman and Martin, 1999).**

	Population ('000)			Households ('000)		
	2001	2006	2011	2001	2006	2011
<i>Natural Growth Only</i>	1,463	1,511	1,545	491	544	604
<i>Natural Increase plus low in-migration</i>	1,471	1,530	1,578	495	554	622
<i>Natural Increase plus medium in-migration</i>	1,481	1,552	1,611	501	566	640
<i>Natural Increase plus high in-migration</i>	1,491	1,575	1,650	507	579	661

From a transportation point of view the strategy for the Metropolitan Area is to:

- Consolidate development within the GDA and increase the overall density; and,
- Thereby provide the basis for a better public transport system and facilitate a shift away from the private car and towards public transport.

This will lead to a more compact urban area and thereby reduce the overall demand for travel provided that demand management measures (such as peak hour road use pricing and parking restrictions) are introduced and implemented (Brady, Shipman and Martin et al., 1999).

The strategy for the Hinterland Area (the rest of the GDA) is to develop a number of towns, each independent in terms of employment, shopping and social facilities. These towns should be self-sufficient, as far as is possible, with little or no commuting to the Metropolitan Area. The spread of development serving the Metropolitan Area and generating a lot of commuting traffic *“is neither environmentally sustainable or economic and should be restricted using demand management techniques”* (Brady, Shipman and Martin et al., 1999).

The SPG recommend that demand management measures should be incorporated into all strategic planning and implementation by the relevant authorities to:

- Reduce the overall growth in the demand for travel;
- Dissuade private car commuting; and,
- Encourage the use of public transport.

The SPG recognise the importance of the private car for transport in the Hinterland Area and that improvement to the road network is required to facilitate this use. However, commuting from the Hinterland Area to the Metropolitan Area should be transferred from the private car to public transport through demand management policy and decisions and through measures that dissuade commuting by private car such as road user pricing.

### 2.6.2 *An integrated transportation strategy for the Greater Dublin Area*

‘A Platform for Change: Final Report’ was published by the DTO in 2001 and describes an integrated transportation strategy for the GDA 2000 – 2016. The report was prepared to support and compliment the strategic land use and transportation planning framework described in the SPG for the GDA, published in 1999 and summarised in Section 2.6.2 above. As such, the SPG are the basis upon which the DTO Strategy rests.

The DTO Strategy has a number of specific objectives for transportation in the GDA, which are summarised here:

- Improve accessibility by reducing the demand for private transport, reducing the need for private car commuting by improving the quality of public transport and by reducing travel times and congestion on the road network to 1991 levels when the average speed in the morning peak hour was 22kph;
- Sustain economic development, foster sustainable land use and transportation development and consolidate existing economic activity;
- Promote the implementation of the SPG for the GDA by consolidating growth within the Metropolitan Region and promoting the self-sufficiency of the 'Development Centres' in the Hinterland Area, see Figure 2.2; and,
- Ensure that legislative, institutional and administrative structures optimise implementation, make efficient and cost effective use of resources and meet sustainable transport needs.

The DTO Strategy to achieve these objectives is an integrated one with two main elements. The Strategy will only be effective if both of these elements are implemented together and in a coherent manner:

- **Demand management** to reduce the growth in the demand for travel while not harming existing or future economic development and encouraging the transfer of trips from the private car to more sustainable modes of transport such as public transport, cycling and walking.
- **Infrastructure and service improvements** to increase the capacity of the transportation network including a substantial increase in the capacity of the public transport network, some road construction and traffic management.



### 2.6.2.1 Infrastructure and service improvements

The public transport elements of the Strategy are expected to provide for 300,000 of the 488,000 trips in the predicted morning peak period in 2016, compared to 70,000 in 2001 (DTO, 2001). This integrated public transport network will incorporate improvements to DART and suburban rail, LUAS light rail system, METRO, an expanded bus network and measures to improve the integration and attractiveness of the public



**Figure 2.10: Map of the existing strategic road network (black) and the proposed strategic road network of the DTO Strategy (DTO, 2001).**

transport network. A detailed map of the completed public transport network infrastructure in 2016 can be seen in Figure 2.11.

A number of national road projects are included in the DTO Strategy for the GDA, see Figure 2.11. These include upgrading orbital routes around Dublin (M50, the Dublin



safety on school trips;

- Local traffic signal control centres to manage traffic, detect faults, observe junctions and communicate problems on the network. The overall objective being to reduce rat-running, pedestrian delays and congestion while increasing bus speeds and improving safety;
- A Regional Freight Study to identify appropriate routes and restrictions for freight and the development of traffic cells and traffic calming; and,
- A motorway control centre for the GDA, improved enforcement systems, planned road works to minimise traffic delay and information systems so people can make informed travel choices.

### 2.6.2.2 Demand management

The SPG states that the success of the land use strategy for the GDA is dependent on the introduction and implementation of demand management measures, since physical planning policies and traffic management can only partially reduce the demand for travel. It has also been accepted by the DTO that the planned infrastructural and service improvements, described earlier, are not capable of achieving the overall objectives of the DTO Strategy and the SPG.

In the Metropolitan Area, see Figure 2.1, the objective is to reduce morning peak hour (08:00-09:00) car journeys forecast for 2016 to approximately 1997 levels (approximately 127,000 car journeys in the morning peak hour in the Metropolitan Area). Within the hinterland of the Metropolitan Region measures should be designed to achieve sustainable modal split and trip distributions for travel within the “*development centres*” identified in the SPG (DTO, 2000) and other urban regions.

A demand management strategy is, therefore, a fundamental element of the overall DTO Strategy. However, a comprehensive study, the Demand Management Study, is necessary to develop this strategy whose primary aims will be to:

- Reduce the growth in overall travel by motorised modes of travel in the Greater Dublin Area;
- Effect further modal transfer from private car to public transport modes over and above that achievable through the infrastructure and service enhancement measures described in *A Platform for Change*;
- Achieve a good level of service on the road network for essential road users; and to,
- Encourage more sustainable trip distributions and modal split throughout the Greater Dublin Area.

The consultant will examine the demand management study tasks laid out below and develop potential demand management measures under these categories:

- Economic or fiscal instruments including vehicle charges, fuel charges and taxes; public transport fares; road pricing and/or congestion charging; and, parking charges including charging for workplace parking;
- Land use policies including advice relating to the location, scale and mix of development; parking standards; development densities and other key development principles to facilitate sustainable travel; and,
- Management and control of public parking, on-street and public off-street parking control and management plus park and ride.

Based on the results from above, a set of demand management measures for the Greater Dublin Area will be recommended, which should meet the objectives listed. The study commenced in September 2002 and is due for completion in January 2004.

### 2.6.3 *Likely benefits to the Greater Dublin Area*

There will be a number of positive impacts associated with implementing the overall DTO Strategy by 2016:

- The transportation system will have adequate capacity for all users during the morning peak period;
- Most people within the GDA will live within a 10 minute walk of the integrated public transport system, including the DART/suburban rail, METRO, LUAS, and a Quality Bus Network, where it will be possible in most cases for journeys to be completed with, at most, one interchange and using only one ticket;

**Table 2.11: Average 2016 AM peak period journey times in the Greater Dublin Area (DTO, 2001).**

	Without the Strategy	With the Strategy
Bus	47	29
Rail	38	18
Car	94	57
Weighted average	76	34

- As a result of the reduction in congestion the average journey times will reduce dramatically with bus, rail and private car average journey times being 37%, 53% and 39%, respectively, less than if the Strategy were not implemented, see Table 2.11;
- With less private cars on the road network, it is predicted that public transport will have 63% of the total transportation market in the GDA and 85% of the market for city centre destinations. This compares favourably to 35% and 72%, respectively, if the Strategy were not implemented, see Table 2.12;
- Some other benefits include reductions of 41% in energy consumed, 34% in emissions and 35% in accidents;

- Access to the airport and the city centre will improve by 43% and 23%, respectively, in the morning peak period; and,

**Table 2.12: Predicted transportation mode split in the morning peak period (DTO, 2001).**

	1991		1997		2016	
Bus	44,000	26%	47,000	19%	69,000	14%
Rail	18,000	10%	21,000	8%	239,000	49%
Car	110,000	64%	181,000	73%	180,000	37%
Total	172,000	100%	249,000	100%	488,000	100%

- The urban areas within the GDA will become more attractive to business, have better employment potential and capacity and become focal points for leisure and retail activities.

According to the DTO, morning peak commuters using their car rather than bus or rail to travel has fallen from 72% in 1997 to 70% in 2002 while the numbers using Dublin Bus have increased by 40%, which represents the first time that bus travel has increased its market share since the DTO started surveying. This increase is due primarily to progress in the completion of the QBCs.

When the DTO produced their traffic model in 1997 they forecast the number of morning peak commuters at 488,000. The current figure is now over 460,000, which gives a clear indication of the size of the transportation challenge being faced (DTO, 2002). It is expected that the forecast of 488,000 will be exceeded in the next few years, well before 2016, even considering the expected levelling off of travel demand.

The transportation deficit in the GDA differs from that of other cities that have implemented congestion charging schemes in that there are infrastructural as well as demand management challenges. For instance, the public transport network and

highway infrastructure of the Greater London Area is very good in comparison to the GDA, yet the demand for road space and congestion difficulties were still high enough for it to be viable to implement a road user charging scheme, see Chapter 3, Section 3.6.2.1 for more details. Similarly, Singapore is a crowded city-state with a good infrastructure taking up 11% of the gross land area of the island (Foo, 2000) but it has still needed to control the demand for road space using demand management techniques including road user charging, see Chapter 3, Section 3.6.2.2 for more details.

### ***2.7 Summary of the Development within the Greater Dublin Area***

- The planning system in Ireland depends on the private sector to put forward potential development projects. As such, it can be considered flexible to pressure for or against development and lacks long term strategic values, see Section 2.2.
- Proposed institutional arrangements include the establishment of a strategic land use and transport body and deregulation of the public transport sector to facilitate a transport market, see Section 2.2.
- The new institutional and regulatory measures should help facilitate the National Development Plan 2000-2006 objectives of addressing projected traffic growth through transport infrastructure and demand management measures, see Section 2.2.
- The majority of reports produced before the DTI began in 1988 failed to address the interactions between land use and transportation, concentrating on one or the other, see Section 2.3.
- The DTI began a transport planning process in 1988 permitting regular review and assessment of the impacts of land use and other policies on transportation and vice versa, see Section 2.4.
- In 1992, Phase 2 of the DTI began resulting in the production of the DTI Strategy in 1994 laying out transportation plans for Dublin, see Section 2.4.

- Due to the fact that the DTI did not have any planning powers they produced a number of possible land use development scenarios with different growth levels for the future, see Section 2.4.
- The objectives of the DTI Strategy were to encourage economic regeneration and development and to help maintain and reinforce Dublin City centre as the country's prime commercial and cultural centre, see Section 2.4 and Appendix 1.
- The DTI Strategy measures included: improving public transport and not letting the private car dominate the transport system; emphasising the movement of people and goods; bringing greater equity to the transport system; and, providing an integrated approach to transportation and land use, see Section 2.4.
- The DTI Strategy could never provide a definite long-term solution because it was based on future land use scenarios. A land use strategy would have to be designed to complement it, which was recognised in 1999 with the production of the SPG, see Section 2.4.
- The 4-stage transportation modelling approach was used to produce the DTNM, which uses trip generation and attraction, trip distribution, modal split and trip assignment (using the SATURN traffic simulation and assignment model), see Section 2.4.4.
- Substantial delays to implementation of the DTI Strategy were unavoidable due, in part, to the division of responsibility amongst a number of agencies, see Section 2.4.
- There has been a huge surge in traffic growth since the DTI began, which has added to delays in implementing the DTI Strategy, see Section 2.5.
- Forecasts for population and traffic growth have been very conservative, actual trends have outstripped all expectations, i.e., predicted population for 2001 was



exceeded in 1997 and the level of demand for travel has increased by double what was predicted for 2001 in the 1991 DTI forecast, see Section 2.5.

- The institutional and regulatory inefficiencies; the cumbersome and time consuming planning system; and, the high economic growth since the DTI began, have resulted in delays to infrastructure projects and unrealistic planning targets resulting in a substantial transportation deficit in the GDA, see Section 2.5.
- In 1999, the SPG, a “*framework for integrated land use and transportation for the sustainable development of the GDA up to the year 2011*”, was produced, see Section 2.6, which provided the first co-ordinated settlement plan for the GDA and the starting point for the latest transportation plans.
- From a transportation perspective, the SPG objectives are to consolidate development and increase density in the GDA so as to provide a basis for a better public transport system, and, to facilitate a shift away from the private car to public transport, see Section 2.6.1.
- The SPG recommend that demand management measures should be incorporated into all strategic planning and implementation to reduce traffic demand, dissuade private car commuting and encourage the use of public transport, see Section 2.6.1.
- The DTO 2000-2016 Strategy also advocates an integrated transportation strategy involving demand management and infrastructural and service improvements, see Section 2.6.2.
- It is recognised that a comprehensive study is necessary to develop a demand management strategy to reduce growth in overall travel by private car and to effect further modal transfer from private car to public transport modes.
- Likely benefits of implementation of the DTO Strategy include between 27% and 57% less congestion than if the Strategy was not implemented. Other

benefits include reductions of 41%, 34% and 35% in energy consumption, emissions and accidents, respectively.

- When compared to the situation in other cities that have implemented congestion charging schemes it appears that the GDA needs some form of demand management measures to be implemented due to the dual problems of having insufficient infrastructural highway and public transport networks, and, a fast growing trip demand.

## *Chapter 3*

# *Road User Charging and the Greater Dublin Area: Literature Review.*

### **3.1 Introduction**

Traffic congestion is a serious problem in many urban centres. Increases in population, car ownership and economic development have led to increased numbers of trips, which have in turn put urban road networks under considerable pressure. In general, investment in public transportation has failed to keep pace with these increases. As a result, the public transport network has, for the most part, been unable to provide an attractive alternative with adequate levels of service and consistent journey times. This failure of public transport to provide a viable alternative to the private car has led to increased congestion and environmental damage in urban areas. With limited funding for the supply of additional transportation infrastructure to cater for the increase in trip demand, it appears that the only option is some form of restraint, in other words, some form of demand management, as suggested by Goodwin et al. (1991). This can be achieved through either regulatory traffic restraint or through increased trip demand management, using road pricing mechanisms in order to better utilise the existing urban road network.

While congestion charging, or congestion pricing, is the application of a charge to reduce the demand for trips, and hence reduce congestion, a broader group of policies is more usually known as road pricing. This broader group includes charging for road use through tolls, parking charges or other means, in order to find new revenue sources for transportation investment and, due to the failure of alternative, more conservative, policies, to significantly reduce the growth of traffic congestion.

The first real comprehensive contribution to the application of pricing theory to road pricing practice and implementation was the Smeed Report, published in 1964 by the UK Ministry of Transport. Over a decade later, in 1975, the first road user pricing scheme was implemented in Singapore. Singapore's Area Licensing Scheme (ALS) was also specifically designed to reduce traffic congestion. There have been other road user pricing implementations since then including the Norwegian toll rings around Oslo, Trondheim and Bergen, and, most recently, a city centre cordon scheme has been implemented in Central London. A number of other trials and studies have taken place

in Hong Kong (1985), the Randstad Region of the Netherlands (1988), Stockholm in Sweden (1991) and in the United Kingdom. Local authorities in the UK, including Leeds, Cambridge, Edinburgh, Bristol and Leicester have taken an interest in road user pricing schemes since new powers were given by the government to allow road pricing to be used to raise revenues for investment (UK DETR, 1998a).

Successful implementations of urban road pricing schemes have, however, been limited to Singapore, Norway and London while single facility congestion pricing has been implemented in France and the USA, in particular. A number of other places are considering the use of road pricing schemes to achieve a range of objectives but have come up against a number of problems, the largest obstacle of which remains the problem of securing political and social acceptance (Jaensirisak, 2002).

This chapter presents a literature review of road user pricing. Section 3.2 describes the development of road user charging, in particular the theory associated with traffic flow and congestion economics and congestion charging itself. Section 3.3 briefly describes the general principles of road user charging design and implementation. Section 3.4 looks at the issues associated with the political and social acceptability of road user charging and ways of increasing the level of acceptability from the general public's point of view. Section 3.5 examines the effectiveness of road user charging using information from pre and post implementation studies. Section 3.6 follows on from this with a brief examination of previous studies in the area and a summary description of existing road user charging implementations. Section 3.7 examines the price elasticities of user demand associated with pricing measures drawing conclusions based on international experience for this research. Finally, Section 3.9 and 3.10 examine demand management in the GDA and conclude the chapter, respectively.

## ***3.2 Development of road user charging***

### *3.2.1 Introduction*

Road user pricing is a transport policy that charges road users for using their vehicles on certain roads or road networks. Jones and Hervik (1992) give two definitions of road

user pricing, one in terms of transport engineering and planning, and the other in terms of economics.

- (1) Engineers and planners define road pricing as the application of a charge for the use of a road. Road pricing can be used as a tool to achieve a number of objectives including the reduction of traffic congestion, the reduction of environmental damage and pollution, and for generating revenue to finance road developments and public transport infrastructure and services.
- (2) Within the field of economics road pricing is defined as the application of a charge equal to the difference between the social marginal cost and the marginal private cost of a journey in order to internalise the externalities associated with the journey, such as traffic congestion, environmental pollution, traffic accidents, etc.

The term *road pricing* is an ambiguous term but is generally used to describe the application of a charge, or toll, for the use of road infrastructure (Thompson, 1990). The Smeed Report (UK Ministry of Transport, 1964) applies the term to both direct and indirect charging of road users. Direct and indirect charging are described by Lewis (1993):

- Indirect methods of charging relate to ownership and usage of a vehicle. Fixed charges such as purchase tax and license fees are usually applied while variable charges such as taxes on tyres, fuel, oil, etc. vary. Other charges for usage of a vehicle include parking tax and area license charges; and,
- Direct charging involves charging specifically for the actual time spent travelling or the distance travelled on the road network. In this case the economic optimum can be exceeded and is used for demand management.

Other terms, such as *road user charging* and *congestion charging* in the UK, *congestion pricing* in the USA (Giuliano, 1992), and terms such as *road tolling*, *value pricing* and *variable pricing* have been used to describe different types of road pricing initiatives.

See the TCRP Report (2003), Chapter 14, for a review of *value pricing* from a U.S.A. perspective.

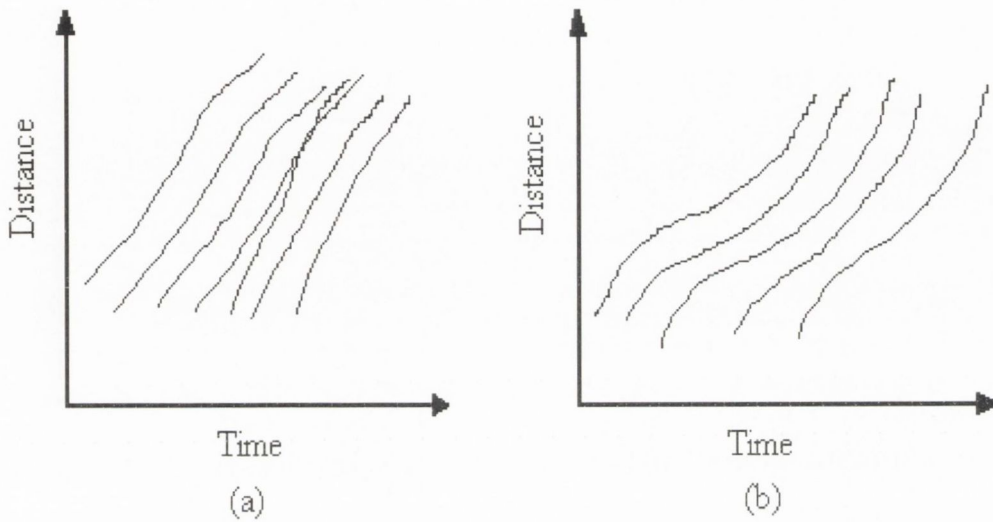
*Road user charging* and *congestion charging* have, in general, been used to describe direct charging schemes while the term *congestion pricing* is only used to describe the pricing objective of reducing traffic congestion. This is done by charging in direct proportion to the levels of traffic congestion encountered by the user (Gomez-Ibanez and Small, 1994). The Transportation Research Board (1994), suggest that congestion pricing be applied during peak periods using direct charging methods. *Variable pricing* has been used to describe road pricing schemes that have variable charges and/or variable periods of charge to encourage a shift of demand to off-peak periods or other modes.

When the charge associated with a road pricing scheme is known as a road toll, this defines the application of a fee by government or private investors upon categories of users of a road infrastructure for one or more purposes, including financing, building, maintaining, operating and improving the infrastructure and associated services. Traffic demand management is not the primary goal of such schemes. In the USA the term *congestion pricing* has been replaced by the term *value pricing* (Orski, 1998), which was introduced in 1995 by the operators of the SR91 Express Lane Scheme in Southern California as part of a marketing plan to make the scheme more acceptable.

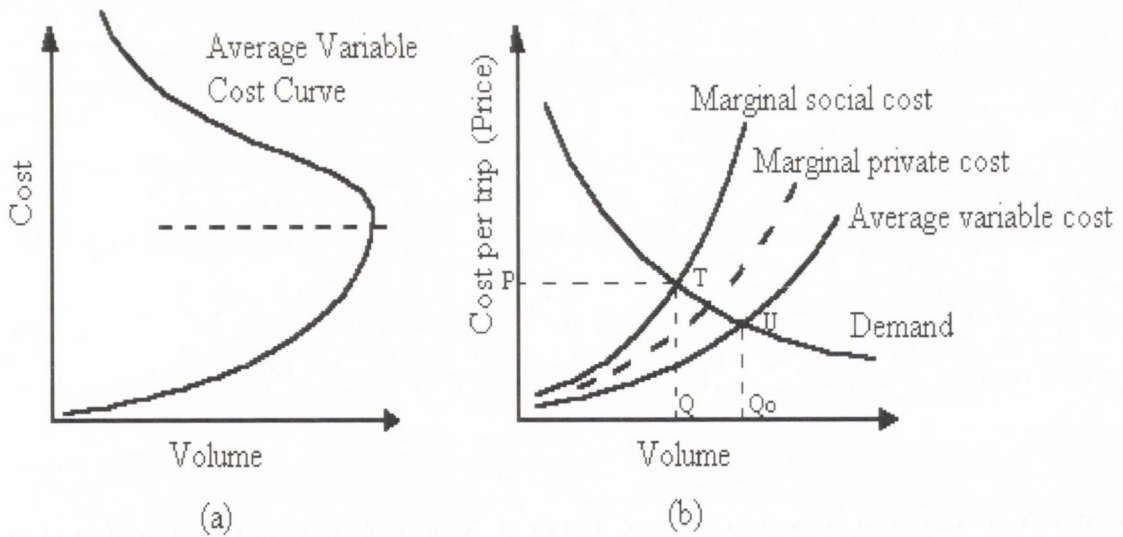
### 3.2.2 *Traffic flow and congestion economics*

Congestion generally occurs as a result of an increase in traffic flow. The volume of traffic consists of vehicles whose drivers wish to minimise their own journey time. The speed and flow of traffic is, therefore, entirely dependent on the behaviour of each vehicle's driver.

On an unobstructed straight road free flow exists where traffic volumes are low and faster drivers are able to overtake slower drivers if they wish to do so. Hence, they are not adversely effected by other drivers' behaviour, see Figure 3.1 (a). This is termed stable flow. Where traffic volume is high each vehicle's speed is dependent on that of



**Figure 3.1: Distance versus time of vehicle trajectories under (a) free flow and (b) unstable flow during a shock wave effect (Lewis, 1993).**



**Figure 3.2: Cost versus volume showing (a) the average variable cost according to volume and (b) the supply contrasted with demand in the lower part of the curve (Lewis, 1993).**

the vehicle in front. Each driver must react, i.e., accelerate or decelerate, depending on the behaviour of the driver in front, establishing a shock wave effect back through the traffic flow, see Figure 3.1 (b). This is termed unstable flow. As the density of traffic rises, speed and the traffic volume decrease. If the road is near capacity the shock wave



effect may result in traffic coming to a halt, resulting in forced flow. As the queue forms and congestion increases, each additional vehicle contributes to the congestion, thereby imposing a delay, and hence a cost, on other drivers (Walters, 1961; UK Ministry of Transport, 1964).

To determine the economic effects of congestion a value of time to travellers must be established. The external cost of congestion is found by multiplying the average value of time by the delay. When the vehicle operating costs and the variable road maintenance costs are added to this the overall cost of congestion can be shown to behave like the curve in Figure 3.2 (a), which represents the average variable cost (AVC) increasing slowly under free flow conditions and asymptotically under forced flow conditions. The AVC curve climbs upward because of the negative interactions that occur before the traffic reaches maximum capacity,  $Q_{max}$ . It is variable in the sense that as traffic flow is increased, congestion delay sets in rapidly well below the maximum capacity,  $Q_{max}$ . This gives us a representation of supply (cost). After the basic 'engineering' maximum capacity,  $Q_{max}$ , is reached the AVC curve becomes an 'inverse supply' curve (Hau, 1992a). Hau (1992a) also notes that the standard supply curve does not exist in the context of roads.

The supply curve can be made congruent with the demand side when the demand curve is specified to depend on the travel cost (price) facing a traveller for a single trip, see Figure 3.2(b). When the demand curve intersects the AVC curve, point U, then a stable equilibrium is said to exist at  $Q_0$  (Hau, 1992a). Theoretically, whenever the AVC curve rises the marginal cost curve lies above it. The marginal cost curves show the private cost of any additional trip as perceived by the driver, and, the higher marginal social cost of the trip representing the cost of delays and other external costs to other drivers, which are imposed by the additional trip. The difference in the marginal private cost for each trip as perceived by the driver, and the social cost that trip imposes in terms of delays on other drivers, represents the marginal (external) congestion cost associated with the trip. Since each driver chooses to drive or not based on the AVC curve, the decision curve, the resulting external congestion cost imposed on other drivers is totally ignored (Hau, 1992a). Therefore, the traveller will not reduce the number of trips that

are made unless this external cost is quantified in the mind of the driver. However, if the driver is charged this difference, thus internalising the externalities associated with each trip, then a more balanced demand will result (Lewis, 1993). In figure 3.2(b) the optimal point, T, at which the social marginal cost curve intersects the demand curve gives the point, Q, being the optimal output in the sense that the total cost of a trip (the marginal social cost including both external costs, operating costs of a vehicle and variable maintenance costs of a road) is equated to the price, or cost.

Road user charging has been widely suggested by transport planners and economists as a solution to the problem of congestion. According to Lewis (1993), a driver's perception of their real cost of travel is key to their behaviour. However, drivers do not perceive these costs and hence, do not take them into account when making the decision to take a trip. Therefore, it is important to devise a mechanism to allow the driver to clearly perceive the real costs of the trip and to charge the external cost, either in terms of delay or marginal social cost, imposed on other vehicles to the driver (Vickrey, 1955).

### 3.2.3 Congestion charging

The theoretical and empirical relationships governing road congestion in urban areas are well founded in the economics and transportation engineering literature (Pigou, 1920; Walters, 1961; Hau, 1992a). Consequently, most urban economists and an increasing number of policy makers agree that the best policy to deal with the problem would be to introduce some form of congestion charging (Small and Gomez-Ibanez, 1998). For instance, the Dublin Transportation Office (DTO, 2000) concluded that some form of demand management would have to be used in conjunction with infrastructural improvements in order to improve the traffic congestion problems in the Greater Dublin Area (GDA). From a European perspective, the High Level Group on Transport Infrastructure Charging (HLGTIC, 1999), convened by the European Commission, concluded that there is a need for an EU level approach to infrastructure charging that includes all the major transport modes. The underlying objectives being to create

efficient transport systems in Europe in order to strengthen Europe's economy and contribute to the 'greater sustainability' of the transport system (HLGTIC, 1999).

Pigou (1920) and Knight (1924) were the first to study the economics of road user pricing in specific detail. They used an example with two roads, arguing that through the application of a toll on a congested road the total travel time would be reduced, and, that the road space would be used more efficiently leading to a net benefit to society. Walters (1954) continued this train of thought suggesting that congestion can be reduced by bringing the marginal private cost of running a vehicle closer to the marginal social cost. Vickrey (1955) also suggested that marginal costs should be considered in price schemes to achieve greater utilisation. However, in a later publication Vickrey concluded that:

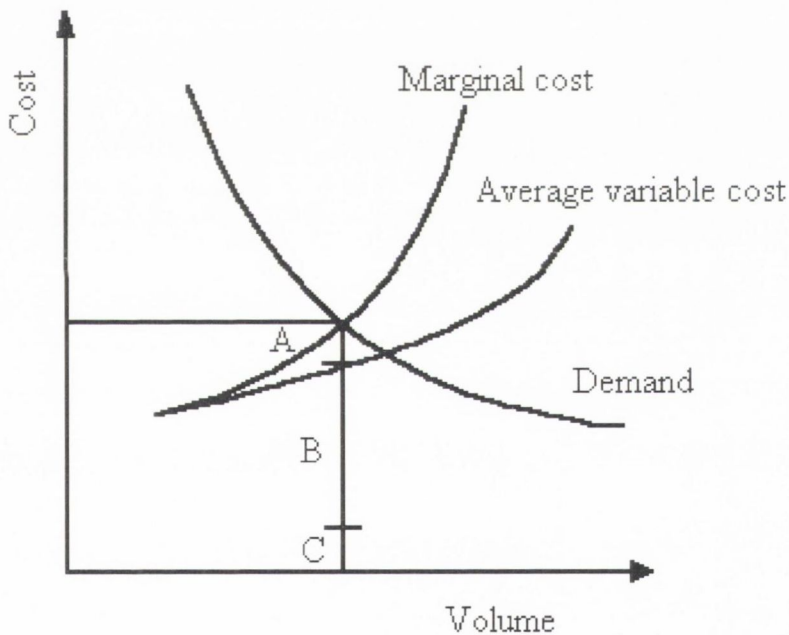
*"In no other major area are pricing practices so irrational, so out of date, and so conducive to waste as in urban transportation."* (Vickrey, 1963)

In 1964 the Smeed Report became a watershed in the study of road user pricing theory and practice and a number of economic studies have taken place since its publication. Some of these include those of Vickrey (1969), Newbery (1990), Small (1992), Verhoef (1996, 1999), and Hau (1992a, 1992b, 1998). The more recent studies have concentrated on the specification of demand and supply for congested networks. However, researchers disagree on the analysis of choice and debate continues.

The flow analysis, see Hau (1992a) and Verhoef (1999), uses a static model with the flow based approach to describe the relationship between the demand function and the supply function, and hence, give a representation of the transportation network performance, see Figure 3.2(b). May et al. (2000) assessed behavioural responses such as route choice and time of travel by measuring the costs to users using individual vehicle tracking across the network instead of during a set time period. Using micro-simulations the issue was discussed further in a UK Department of the Environment, Transport and the Environs project entitled 'Analysis of Congested Networks' (UK DETR, 2001), carried out by the Institute of Transport Studies (ITS, Leeds), Transport Operations Research Group (TORG, Newcastle upon Tyne), John Bates' Services and

the Transport Research Laboratory (TRL). It was found that using performance curves to estimate supply curves was unreliable in that the levels of congestion observed where congestion charging would be recommended are overestimated while the effects of congestion charging schemes are underestimated.

Figure 3.3 shows the economic components involved in road user pricing. In examination of this figure it is noted that a driver currently makes the decision to undertake a journey based on the perceived average variable cost of that trip. This includes vehicle operating costs, time costs and road maintenance costs paid through



**Figure 3.3: Graph of the economic components involved in road user pricing (Lewis, 1993).**

taxes. External costs associated with the trip are not considered in the decision process. In terms of road user pricing economics, the equilibrium point occurs where the demand intersects the marginal cost curve. At this point the vertical difference (A) between the average variable cost curve and the marginal cost curve represents the external congestion cost resulting from an additional trip. The difference (B) refers to the user operating cost and time borne directly by the user and (C) refers to the road maintenance cost. Verhoef (2003) looks at key implementation issues associated with the introduction of marginal cost pricing in transportation. From a demand management

perspective this economic optimum pricing charge may be exceeded in order to restrain or reduce traffic volumes. At this point road user pricing becomes congestion charging where the charge is designed specifically to reduce trip demand and not just to internalise external costs.

A successful urban road user charging system must address the viewpoints of the different groups that are affected by it (Lewis, 1993). While the social aim is to ensure that the social welfare function is maximised, i.e., that there is an overall net benefit to society, it is recognised that costs and benefits to each group will vary in terms of time and money. Users with high values of time may be better off at the expense of others. The London Congestion Charging Research Programme into road pricing has identified these Impacted Behaviour Groups (IBGs) and their respondents (MVA, 1995), see Table 3.1. Table 3.1 shows a list of respondents who are impacted in different ways depending on their lifestyle. For instance, a respondent falling into the category group ‘traveller’ will most likely describe strong behavioural impacts to their mobility.

The goal is to ensure a positive benefit/cost ratio for any particular group (Lewis, 1993). However, optimising the benefit/cost ratio may not correspond with the optimisation of the revenue/cost ratio. In other words, though the optimal pricing charge may be

**Table 3.1: Impacted behavioural groups (IBGs) and their respondents (Lewis, 1993).**

<b>Group</b>	<b>Respondent</b>
Spatial	Land user
Socio-economic	General public and residents
Mobility	Traveller
Journey purpose	Employee/shopper/recreation
Public service provider	Health/education
Transport provider	Operator
Business	Employer

desirable from an economic standpoint, from a social or practical viewpoint, e.g., congestion charging, deviation from the economic theoretical optimum is more than likely, which inevitably makes it more difficult to ensure a net benefit for every group. This leads to a discussion of ‘first-best’ and ‘second-best’ congestion charging.

‘First-best’ congestion charging is a theoretical ideal. ‘First-best’ congestion charging would be to charge each road user a fee that is equivalent to the increased costs the users’ presence on the road implies for all other road users determined at a precise level. As such, the charge level should theoretically vary not only with the time of day and the type of road but with the specific time values of the road users that happen to be present on the road at the same moment. This is currently too complex for practical implementation on a road network. Therefore, any ‘second-best’ congestion charging scheme, referring to any scheme that is not ‘first-best’, that might be implemented is simplified as much as is necessary and be predictable to the road user as to the charge level (Mattsson, 2003b). For instance, road user charging should be closely related to the amount of direct usage so that a driver who uses the network more pays more (Lewis, 1993); however, in practical implementations (Singapore and London) caps have been placed on the maximum that a driver can be charged per month, which appears to limit this ‘first-best’ ideal for practical reasons resulting in a ‘second-best’ solution.

### ***3.3 Urban road user charging design and implementation***

#### ***3.3.1 Introduction***

Any road user charging system needs to be equitable, efficient and acceptable to the majority of users (Jones, 1998). The objective of an urban road user charging system should reflect the policy objectives of the governing transportation authority. This, in turn, will be reflected in the authority’s transport planning framework. This planning framework is usually largely concerned with matters associated with trip generation-attraction and trip distribution. However, trip generation-attraction and distribution are generally dependent on a number of factors including topology and land-use,

residential, commercial and industrial location, the existing transportation network, and, other factors that may be out of the control of the transportation authority.

Therefore, the policy objectives of the authority should be clearly defined with respect to road user charging. These objectives may be to reduce congestion, raise revenue for transportation infrastructure, reduce the demand for travel and hence the level of environmental pollution, reduce the demand for the construction of new roads or to act as a rationing mechanism for scarce road space.

### 3.3.2 *Basic design criteria*

One of the most important areas of focus in reducing traffic congestion is the reduction of non-essential vehicle trips in peak periods (Lewis, 1993). A list of basic design criteria has been formulated for the implementation of road user charging schemes in the Smeed Report (UK Ministry of Transport, 1964):

- The majority of users should find the road user charge, and the charging system, acceptable and fair;
- The road user charging scheme should be both easy to understand and to use;
- The system should have a high degree of reliability;
- The road user charging scheme needs to be accurate and the public need to be able to verify prices before they embark on a journey; and,
- The scheme should be enforceable against both fraud and evasion, be it either deliberate or unintentional.

Other criteria have been added since then by Thompson (1990) and Hau (1992b), which have been referenced by Lewis (1993) as follows:

- The charging scheme should be designed to protect individual users' right to privacy while allowing them to check their account and charges levied;

- The scheme should be able to integrate with other services and infrastructural systems, particularly those associated with driver information systems; and,
- Technology should be utilised to achieve a road user charging scheme based on the criteria set out above.

### 3.3.3 *Practical aspects of design and implementation*

From a practical standpoint a number of studies have been produced that look at the main aspects of implementing a road user pricing scheme. These include work by May (1992), Hau (1992b), Lewis (1993), Gomez-Ibanez and Small (1995), Small and Gomez-Ibanez (1998), and, Jones (1998). Once the objectives of the scheme have been ascertained five specific areas of practical importance must be addressed before implementation of the scheme. These have been described by Jones (1998) as:

- The type of traveller or vehicle to be charged;
- The area where the charge is to be applied;
- The time period in which to apply the charge;
- The level of charge; and,
- The characteristic on which the charge is based, i.e., based on the amount of time spent travelling, or on the distance travelled, etc.

Categorising the vehicle types to be charged becomes easier once the objectives of the scheme have been defined. Jones (1998) suggests that some types of traveller or vehicle should be made exempt depending on the objectives of the scheme. For instance, pedestrians and cyclists are always exempt because there are no external costs associated with their movements. If the scheme objective is to reduce pollution then energy efficient cars may be made exempt, while in the case of congestion reduction public transport may be made exempt. Other travellers or vehicles to be considered



include the emergency services, goods vehicles, residents within a charged area, motorcyclists, disabled drivers and high occupancy vehicles.

When considering the charged area there are three categories of implementation size, which have been noted by Decorla-Souza (1993). Once again the scheme objectives determine the scale of the system to be implemented. The most basic definition for the charged area is used in single facility pricing, where a section of road or a bridge represents the charged area, see Section 3.6.5 for an example. The most common type of charged area is used in area-wide pricing, where a cordon is applied to a specific area, usually a city centre or a central business district. This type of scheme has been introduced in Singapore, the Nordic cities of Oslo, Trondheim and Bergen, and, most recently, Central London. The third type of area involves large scale region-wide pricing, where a large urban regional area is covered by the scheme; for example the Ranstad Region of the Netherlands, see Section 3.6.4.

The duration and time of day of the charged period is also dependent on the scheme objectives. For instance, in the case of a required congestion reduction the charged period may be one or two hours during the morning peak period or a blanket congestion charge during the working week like in the London congestion charging scheme, see Section 3.6.2. However, where revenue collection is a priority a 24-hour charged period may be used in conjunction with a lower charge like that in Oslo. The level of charge is also closely related to the scheme objectives. A higher charge is generally used for reducing traffic volumes and pollution whereas a lower charge would be used where revenue collection is a priority and the controlling organisation does not want to price users off the road, for example, the Nordic cities of Oslo, Trondheim and Bergen (Small and Gomez-Ibanez, 1998). The charge level can also vary depending on the type of user or vehicle, by time of day, by area, or depending on the direction of traffic.

Jones (1998) describes two broad bases upon which road user charging schemes can be described; area-based and point-based. Within the point-based group there are cordon and cellular oriented schemes (MVA, 1995). Cordon schemes involve one or more boundaries and may also use a number of screen lines. Cellular schemes are similar but

use a number of cells over a wider area. Both types charge when a boundary is crossed, though ceilings can be put in place for the maximum amount one person can be charged per day, week or month.

There are many types of area-based charging schemes including those which are congestion oriented, time oriented, distance oriented, externality oriented or make use of supplementary licensing. Congestion schemes apply a charge to users travelling during congested periods in the designated charge area. When congestion dissipates then, ideally, so should the charge. This theoretically ideal congestion charging scheme, where charge is applied directly in relation to the levels of congestion experienced, was attempted in Cambridge (Oldbridge, 1995), see Section 3.6.2. This type of scheme can also use a charge based on the delays experienced by users but this may induce unsafe driving behaviour (Bonsall and Palmer, 1997).

Time-based charging schemes involve charging users for the amount of time spent travelling within the charge area. While this may seem like quite a fair method of charging it may induce unsafe driving practice and lead to an increase in accidents, although there are no studies to prove this. This contrasts with distance-based charging schemes, which are similar but more predictable because of users' route choice and, according to Bonsall and Palmer's (1997) study using a driving simulator, are far less likely to encourage unsafe driving behaviour. It should be noted that no full scale real studies have been undertaken to prove this conclusively.

Singapore's Area Licence Scheme (ALS), from 1975 to 1998, has been the best example of the use of a supplementary licence scheme to date. May (1975) defines such a scheme as the purchase of a licence to be displayed on any vehicle used within a charged area. This scheme was replaced in 1998 by an ERP system (Menon and Chin, 2000), which has also been introduced in Central London (GOL, 2000).

### **3.4 Political and social acceptability of road user charging**

#### *3.4.1 Introduction*

To implement an urban road user charging system; political will, social acceptance, the willingness of the public to pay the charge, budgetary constraints and the availability of alternatives are fundamental to the success of the scheme (PRIMA, 2000). The lack of public acceptance is probably the largest obstacle to implementation of road user charging (Jones, 1998). The public must recognise an overall net benefit to their own independent interests. The problem has been well formulated by Zettel and Carll (1964). Once an urban road user charging system has been introduced people fall into three broad groups:

- (1) Individuals who are 'tolled' must pay for a commodity that used to be free (Giuliano, 1992; Small, 1992).
- (2) Individuals who are 'tolled off' must use less desirable modes of transport, routes or times of travel.
- (3) Individuals who are 'untolled', i.e., public transport users, are worse off because of increased congestion of the public transport network.

Hence, almost everyone, except those with a very high value of time, is worse off. It is therefore necessary to implement a policy that makes peoples lives better off, improves the environment and makes the economy more efficient (Goodwin, 1997). To do this it is imperative that revenues gained through road pricing are shown to be used in such a way as to increase public benefit. Making the public more aware of the positive benefits of a road pricing scheme can be done through a number of ways including investment in road developments, the environment and public transport infrastructure and services, and, by reducing vehicle related taxes. When the public are made more aware of the likely benefits then they are more willing to pay (Transportation Research Board, 1994).

### 3.4.2 *Acceptability of road user charging*

Each individual of the public is a member of a different class or group of people. Once their behaviour has been affected they become part of one of the Impacted Behavioural Groups (IBGs) or interest groups (Small, 1992). These groups carry more relative weight than individuals and can lobby on behalf of their members to exert pressure. Any policy initiated by a governing, or transportation, authority must, therefore, attract the support of the strongest, or largest, IBGs while trying to avoid alienating minority groups. See Table 3.2 for a typical breakdown of interested groups.

In 1991 surveys in the United Kingdom showed increasing public support for measures to overcome traffic congestion, which included road pricing, provided that other elements of traffic restraint and an enhancement of public transport alternatives were included (Jones, 1991). The most common general requirements were:

- Restriction of cars in inner cities;
- Improvement of public transport alternatives;
- Enforcement of parking controls;
- Improvement of road links and intersections, and, traffic management;
- Introduction of traffic calming measures; and,
- Improvement of pedestrian and cycle routes.

Several attitudinal surveys towards road user charging have been carried out in the UK over the last decade, see Table 3.3. On examination of these results large differences are noted. There are a number of reasons for this. In London the results are far more positive than for the UK nation-wide surveys. This could be due to the fact that there is a higher percentage of non-car users in London. GOL (2000) reported that 30% of car users thought road pricing to be a positive idea, whereas 54% of the general sample agreed with the concept of road pricing. Another possible explanation is simply that congestion is far more of a problem in London than in other smaller urban centres and

**Table 3.2: Interest groups (Small, 1992).**

<b>Industries</b>	<b>User groups</b>	<b>Environmental groups</b>	<b>Residents</b>
Motor	Motorists	Walkers/cyclists	Pedestrians
Freight	Truckers	Anti-pollution	Recreation
Public Transport	Passengers	Rail travellers	Shoppers
Road building	Contractors	Conservationists	Home owners

**Table 3.3: Acceptability of road user pricing (Partial source: Jaensirisak, 2002)**

<b>Case study</b>	<b>Source</b>	<b>Year of survey</b>	<b>Results</b>
<b>UK Nationwide</b>	Jones (1991)	1991	30% support
	Taylor and Brook (1998)	1993, 1995, 1996	18% support of pricing in city centre (1993) 25% support of pricing in city centre (1995) 30% support for £2 charge for entering city centre at peak times (1996)
	CfIT/MORI (2000)	2000	27% support
	CfIT/MORI (2001)	2001	37% support
<b>London</b>	NEDO (1991)	1991	43% acceptance
	Halcrow Fox and Associates (1992)	1992	37% acceptance of car users
	London First (1999)	1999	76% support (£5 daily car charge between 7am and 7pm inside Inner Ring Road with resident discounts)
	GOL (2000)	1999	53% of respondents positive 30% of car users positive (£5 daily car charge in Central London and £2.50 for Inner London)
<b>Oslo</b>	Oscar Faber et al. (1998)	1990	33% positive, 66% negative at opening of toll ring 50-60% using automatic toll lanes in 1998
<b>Bergen</b>	Oscar Faber et al. (1998)	1986	54% negative, 13% positive before ring implementation >50% positive within one year of operation
<b>Trondheim</b>	Oscar Faber et al. (1998)	1991 1996	72% negative, 8% positive 6 months prior to opening ring 48% negative, 19% positive 2 months after opening (1991) 36% negative, 32% positive, 27% indifferent (1996)
<b>Cambridge</b>	Thorpe et al. (2000)	1994	34% acceptance of city centre car charge (73% travelled into the city centre once a week or less)
<b>Leeds</b>	Schlag and Schade (2000)	Not Known	8-16% support for different pricing schemes: distance pricing, congestion pricing and cordon pricing

the public don't see too many other alternatives to address the congestion problem other than road pricing. These differences may also be accounted for where different road user charging scheme characteristics were presented. However, the main difference in results is most likely associated with the presentation of different charging schemes themselves, which tend to have considerably different levels of social and political acceptance. With regard to the different types of road user charging schemes, results referenced by Glazer et al. (2001) appear to confirm a negative response from the public to differential road user pricing such as distance-based, time-based and delay-based charging when compared to a fixed charge type scheme. Taking London as an example it is noted that of the two surveys conducted in the year 1999, the London First (1999) survey resulted in a 76% support while the GOL (2000) survey resulted in a much lower 53% support. The GOL used a daily charge of £5stg for driving in Central London and £2.50stg for driving in Inner London. London First (1999) presented a scheme where a charge of £5stg per day between 7am and 7pm would be charged inside the Inner Ring Road and a discount of £104stg per annum would be afforded to the residents within the cordon. It appears that the London First scheme was far more acceptable to the general public with 23% more support. See Section 3.4.3 for more detail on the effects of the different types of road user charging on acceptability.

Luk and Chung (1997) have found that in other countries the range of acceptability usually lies somewhere between 15% and 50% acceptance, unless the public have practical experience of a successful charging scheme, in which case there is an increase in acceptance, for example, in Singapore where the level of acceptance has risen to above 50%. The Oslo experience has also demonstrated that acceptance increases through successful implementation. In 1989 acceptance was at a level of 28%, which increased to 40% by 1995 (Odeck and Brathen, 1997) and again to 46% in 1998 (Harsman et al., 2000).

### *3.4.3 Increasing acceptability of road user charging*

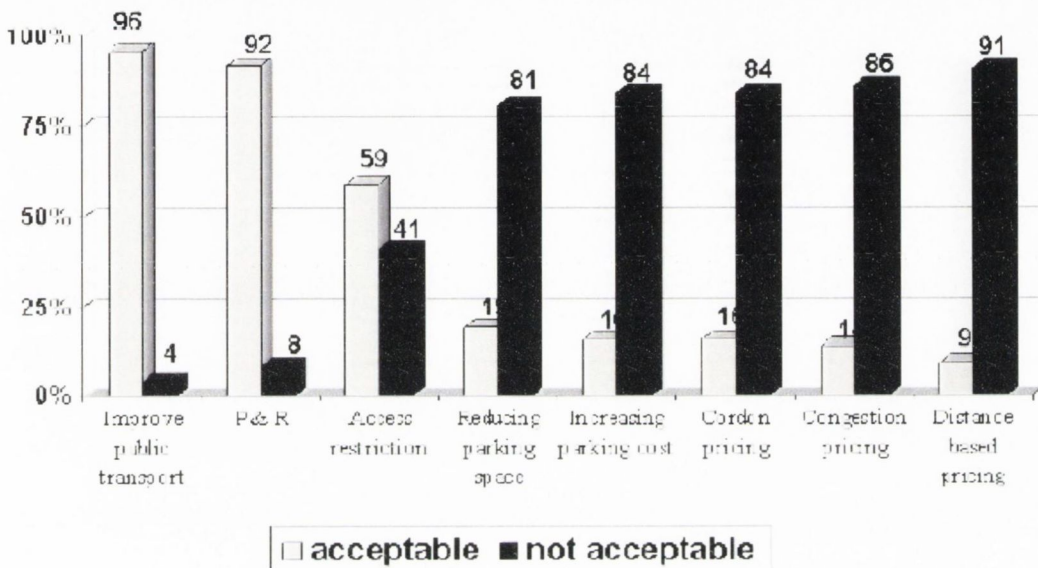
Jaensirisak (2002) describes a number of factors that affect acceptability. The first relates to the stated benefits of the scheme, defining the scheme objectives while

meeting public concerns. A number of studies have addressed this including Sheldon et al.'s (1993) London congestion charging study; Odeck and Brathen's (1997) paper on the Oslo toll ring; and, the TransPrice research project (Schlag and Schade, 2000; TransPrice, 1999).

The stated benefits of a scheme include congestion reduction, or an increase in time savings, environmental benefits and revenue collection. However, in the case of improved journey times, Jaensirisak et al. (2003) and Harrington et al. (2001) have shown that time savings do not have a significant impact in improving the acceptability of road user charging schemes. Other studies have suggested that using collected revenues to improve public transport, reduce travel related taxes or improve the environment result in a positive increase of public acceptance (Jones, 1991; Stokes, 1996; CfIT/MORI, 2000; GOL, 2000; Thorpe et al., 2000). For example, in CfIT/MORI's 2001 survey an increase in acceptance of 14% was found when revenues were invested in public transport (CfIT/MORI, 2001). Jaensirisak (2002) has found that using road user charging as a financial tool is more easily accepted by the public than using it for demand management.

The second major factor affecting acceptance are the scheme features themselves, as described above in Section 3.3.3. Schade and Schlag (2003) have found that complex schemes that are time-based or delay based are less likely to be accepted. This is more than likely because they are not easily understood or easily used by the public. Schade (2003) describes results from the European research project TransPrice that show the different levels of acceptability for various demand management measures, see Figure 3.4, with distance based charging being the least acceptable of the road user charging measures surveyed. In a related example Bonsall and Cho (1999) found that schemes where the charge is unknown or is difficult to ascertain are more likely to be unpopular. The methods of charge used by a scheme and the physical characteristics of the scheme have also been found to affect acceptability (Schlag and Schade, 2000; London First, 1999; GOL, 2000). In London it was found that the majority of the public (74%) preferred a £5stg charge for Central London between 7am and 7pm (London First, 1999), compared to the GOL (2000) pricing scheme presentation of a £5stg charge for

Central London and a £2.50stg charge for Inner London during the day, which garnered a 53% approval from the public.



**Figure 3.4: Acceptability of various demand management measures (Schade and Schlag, 2003; Schade, 2003).**

Other factors that affect acceptability include attitudes towards problems in transportation and the perceived effectiveness of any pricing scheme to address these problems (van der Loop and Veling, 1994; Schlag and Toubel, 1997 (cited in: Schade, 2002); Rietveld and Verhoef, 1998), as well as environmental concerns (Nilsson and Kuller, 2000) and issues associated with freedom (Baron, 1995; Jakobsson et al., 2000) and equity (Giuliano, 1992; Transportation Research Board, 1994; Langmyhr, 1997; Litman, 1999). Personal characteristics and constraints such as income, age, education, etc. also have an effect (Stokes and Taylor, 1995; Odeck and Brathen, 1997; Rietveld and Verhoef, 1998; Harrington et al., 2001).

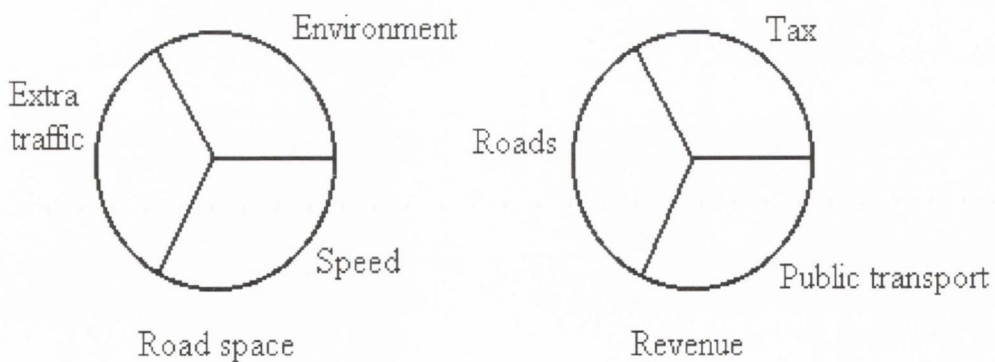
Methods of improving the acceptance of road user charging schemes have been addressed in a number of studies, including Giuliano (1992), May (1992), Sheldon et al. (1993), the TransPrice project (Schlag and Teubel, 1997; TransPrice, 1999; Schlag and Schade, 2000), Jones (1998), Rietveld and Verhoef (1998), the PRIMA project (Gueller,



2000; Harsman et al., 2000), Harsman (2001), and, Jaensirisak (2002). These include clear scheme objectives and an outline of the likely benefits of any scheme to society. The scheme needs to be simple to understand and charges should be quantifiable before the trip is undertaken. This can be achieved through the application of a simple pricing system but also by using a marketing strategy to emphasise the scheme characteristics and by allowing the scheme to be perceived as a good solution to the transportation problems (GOL, 2000).

Revenue redistribution, or allocation, needs to correlate with public opinion, which usually means assigning funds to public transport, environmental improvement, reducing motoring taxes and infrastructural improvement. Revenue redistribution is closely related to equity; the scheme must be perceived as being both fair and just (Langmyhr, 1997). Litman (1999) examines how economic efficiency, equity, external costs and political feasibility can help determine the distribution of road user charging revenue. To address both freedom and equity issues alternatives need to be made available, i.e. better public transport provision for those priced off the road.

As an approach to the redistribution of raised revenues from road user charging Goodwin's 'rule of three', see Figure 3.5, represents an equitable example whereby the



**Figure 3.5: Expenditure of revenue and allocation of road space, Goodwin's 'rule of three' (Goodwin, 1989),**

benefits from road user charging are distributed through the re-allocation of road space and revenues earned (Goodwin, 1989). While it is accepted that anomalies will occur in

any cordon or area based charging system, the main aim is to introduce as equitable a system as possible by being both pragmatic and flexible in the approach to its implementation (Lewis, 1993).

### 3.5 *Effectiveness of road user charging*

#### 3.5.1 *Introduction*

A large body of literature has been produced looking at both design and evaluation of congestion charging. Notable contributions include those from William Vickrey (1955, 1963, 1969), Walters (1961), UK Ministry of Transport (1964), Mohring (1963), May (1975), Gomez-Ibañez and Fauth (1980), and Small (1992). The US National Research Council (Transportation Research Board, 1994) has concentrated on implementation while Grieco and Jones (1994), and Emmerink *et al.* (1995) have evaluated different policy approaches.

The effectiveness of road user charging depends on the objectives of the scheme itself. A user response is the most obvious sign of the effect of road user charging, usually reflecting a reduction in traffic or a switch in mode from car to other modes.

#### 3.5.2 *Studied effects of road user charging*

The behavioural response to road user charging schemes can be measured using a number of different methods, including network modelling, stated preference surveys, attitudinal surveys, field trials and the observed response to practical implementations. Table 3.4 gives an international review of the effects of road user charging studies. These different types of assessment are necessary due to the complex nature of introducing a congestion charging scheme and the expansive changes that result. In practice, the modelling approach used for London by GOL (2000) proved to be the most accurate when compared to actual implementation.

From a modelling perspective, effects from road user charging schemes are expected. In London a number of modelling studies have been conducted, including those by MVA Consultancy (MVA, 1995) and the Government Office of London (GOL, 2000).

The London Congestion Charging Research Programme (MVA, 1995) found that the application of an inbound charge of between £2stg and £8stg from 7am to 7pm would provide between an 8% and 22% reduction in the total vehicle kilometres travelled in the city. At £8stg, 21% of user respondents would switch mode to public transport, with an increase of 16% predicted for inbound bus passengers during the morning peak, and, 3% and 4% increases of suburban rail and underground usage, respectively. In 2000, the Road Charging Options for London (ROCOL) study predicted that car travel to or from Central London would reduce by 20%, and that home-based commuting trips would fall by 30% (GOL, 2000). In the case of London, transport planners have been able to see how accurate these models have been because of the implementation of the Central London congestion charging cordon scheme.

The Central London scheme has now been in place for a period of six months. A three month report has been produced by Transport for London that demonstrates the effectiveness of a city centre congestion charging cordon scheme (Transport for London, 2003). An overall reduction in congestion of 32% has been achieved with a 16% reduction of traffic levels inside the cordoned Central London area. This compares favourably with the 10-15% estimation in the ROCOL report (GOL, 2000). It was also found that 50-70% of the daily reduction, some 75,000–105,000 car movements, have transferred to public transport. A further 20-30% has switched to other modes and 10-20% of the reduction is represented by car journeys, which previously travelled through the zone and are now diverting around it. In Singapore the initial Area Licence Scheme (ALS) was too effective with a 45% reduction of traffic levels, 15-20% higher than targets. The fees were simply set too high (Watson and Holland, 1978; Wilson, 1988; McCarthy and Tay, 1993).

A modelling study has also been carried out in Dublin. It examined a small number of testing scenarios to try to assess the feasibility of road use pricing and other fiscal instruments, such as parking charges, for the management of transportation and traffic demand in the GDA. Though the modelling was not detailed, a 12-25% reduction in vehicle hours was forecast for the city centre using a charge of IR£3 applied at a cordon of approximately the same size as the Oslo toll ring (Oscar Faber et al., 1998).

**Table 3.4: Effectiveness of road user charging assessed through an international review of pre and post implementation studies (Partial source: Jaensirisak, 2002).**

Case study	Source	Method	Charging level	Results
UK Nation wide	Taylor and Brook (1998)	Attitudinal survey	£2 city centre	35% reduction of car use / reliance
London	MVA (1995)	Modelling	£2 / crossing	8% reduction of total vehicle kilometres
			£8 / crossing	22% reduction of total vehicle kilometres 21% increase in public transport usage
	LEX (1999)	Attitudinal survey	£3-£6 per day	30-39% change to public transport
	GOL (2000)	Attitudinal survey	£5 per day	30% not buy a licence (half change to public transport)
		Modelling	£5 per day	20% reduction of car trips in Central London 30% reduction of home-based commuting trips 10-12% reduction of total traffic level
Transport for London (2003)	Experience	£5 per day 7.00-18.30 daily	16% reduction of traffic levels inside Central London 32% overall congestion reduction and of that 50-70% transfer to public transport	
Dublin	O'Mahony et al. (2000)	Field Trial	Average 6.4 Euro per trip	22% reduction from suppression and change to public transport (distance and time-based charging)
	Oscar Faber et al. (1998)	Modelling	IR£3 city centre cordon	12-25% vehicle hour reduction in city centre
IR£3 inner and outer cordon			21% vehicle hour reduction in Dublin Metropolitan Area	
Singapore	New Nation, 1975; Goodwin and Jones, 1989, Menon et al. (1993)	Experience	Area Licensing Scheme	45% reduction, morning peak ALS (1975) 44% reduction inbound traffic with intro of evening ALS (1989) 36% increase in work trips by bus share (1993)
Leeds	Richards and Harrison (1999)	Modelling	£2 peak £1 off peak	11% city centre traffic reduction
	May et al. (1998)	Stated preference	£0.50	25% car use reduction
			£1.00	54% car use reduction
£2.00	90% car use reduction			
Edinburgh	Edinburgh City Council (2001)	Modelling	£1 / crossing	5% increase in public transport usage
			£4 / crossing	15% increase in public transport usage
Leicester	Smith and Burton (1998)	Field Trial	£1.5 tolling 7.45-8.45am	20% car use reduction
			£3 tolling 7.45-8.45am	18-22% car use reduction
			£3 tolling 7.00-10.00am	29% car use reduction

Additionally, the combination of an inner and outer cordon would result in a 21% reduction of vehicle hours throughout the Dublin Metropolitan Region. A field trial conducted by O'Mahony et al. (2000) examined time and distance based charging using an average cost of 6.4 Euro per trip, which resulted in a 22% reduction in congestion. In Leeds MVA Consultants found that a £2stg peak charge combined with a £1stg off-peak charge could reduce city centre traffic by as much as 11% (Richards and Harrison, 1999).

It is noted from Table 3.4 that there are differences in the predicted effects of congestion charging schemes when compared to practical implementation, as discussed above. However, it is also noted that there are differences in the affects when different types of predictive method are used to assess a particular charging scheme. For example, in the case of the Central London congestion charging scheme, a number of different predictive methods were utilised including modelling and attitudinal surveys, see Table 3.4. When the results from these methods are compared to the practical implementation of the congestion charging scheme it has been found that modelling more accurately predicts the likely effects on congestion and trip reduction while attitudinal surveys are better at predicting the likely transfer to other modes of transport. In London, attitudinal surveys predicted that 30-39% (LEX,1999) of those surveyed would change to public transport while the ROCOL Report (GOL, 2000) predicted in its attitudinal survey that of the 30% who would not travel by car, 50% would travel by public transport. Practically, 50-70% of those no longer using their car have transferred to public transport (Transport for London, 2003), which, although higher than the attitudinal surveys predicted, is relatively accurate. From a modelling viewpoint, the most up to date study conducted by GOL (2000) predicted a 10-12% reduction in total traffic level while MVA (1995) predicted a 22% reduction in vehicle kilometres for an £8 charge. Practically, a 32% reduction in congestion has been found, which seems to signify that the scheme is far more effective than was predicted.

Users who respond to the charging schemes usually do so in two ways. Raux et al. (1998), simulating behaviour in an urban road pricing environment, describes these responses as either immediate changes or extraordinary changes. Immediate changes

are postponement or cancellation of trips, route changes or rescheduling of activities for different times. Extraordinary changes include mode switching, changing jobs or relocation. As seems intuitively apparent May et al. (1998) and Bosnall et al. (1998) concluded that users are more likely to respond through immediate changes than extraordinary ones. They also reported an order in which users respond:

- Take a different route;
- Travel before charged period;
- Travel after charged period;
- Use park and ride;
- Use public transport;
- Use car share; and,
- Relocate.

Taylor and Brook (1998) describe a 35% reduction in the use of, or reliance on, cars as a result of a £2stg charge for entering a city, found in the British Social Attitudes surveys of 1996. LEX's (1999) attitudinal survey in London reported that 30% of car commuters would switch mode to public transport if a daily charge of £3stg was implemented, and this would increase to 39% for a £6stg charge. Similarly, the RCOL report (GOL, 2000) found that with a £5stg daily charge for Central London 30% of car users would not buy a licence, and over half of these would switch to other modes of travel. In Singapore, the use of bus services increased from 33% of work trips before ALS to 69% in 1983 (Goodwin and Jones, 1989). Modelling in Leicester and Edinburgh reinforces these findings. Smith and Burton (1998) found that 20-30% would switch from car travel to other modes when tolled between £1.50stg and £3stg in the peak period. In Edinburgh, a city centre cordon charge of £1stg-£4stg would result in an increase in public transport usage of 5-15% respectively. Finally, May et al. (1998) has reported findings from a stated preference survey that show between 25%

and 90% of car users would change travel time or mode of travel if a cordon charge of £0.50stg-£2stg was implemented during the morning peak period in Leeds.

The type of charging system used also affects congestion reduction. May and Milne (2000) modelled four charging systems, using the SATURN traffic simulation and assignment model, based on time travelled, distance travelled, time spent in queues and cordons crossed using the Cambridge road network. It was concluded that, overall, time based charging performed best. However, congestion oriented charging increased network speed most effectively but encouraged re-routing around charged areas resulting in extra distance travelled.

From the individual's perspective the journey purpose also affects car usage and user response. The reduction of car use for work related trips is higher than for other purposes (Oscar Faber et al., 1998; GOL, 2000). Commuting trips are less flexible than other trips such as leisure trips in that they usually finish at the same destination and must be completed by the same time on a daily basis. It has been found that switching mode from the car to public transport is the most likely choice for work related commuting trips (Gärling et al., 2000).

Road user charging has other effects. One of the most important of these is the generation of revenue, which can then be used to address matters of equity, fund improvements to public transport infrastructure and services, reduce vehicle associated taxes and fund other projects. The Central London congestion charging scheme is expected to raise more than £1.3bn sterling over 10 years of operation, which must by law, be re-invested in London's transport infrastructure; see the Greater London Congestion Charging Order produced by Transport for London having exercised powers assigned to it under the Greater London Authority Act 1999 (Transport for London, 2001). This order states that all of the revenues collected through the scheme, approximately £200 million sterling annually, must be spent on transport related projects to make improvements to: the bus network and accessibility, public transport infrastructure and services (for instance, increasing late night public transport), safety

and security, the walking and cycling environment, and, Greater London's public transport rail and bus capacities.

Richards and Harrison's (1999) modelling of an inner city congestion charge in Leeds has led to a prediction of £25 million sterling in net revenue per annum, which would be used in a similar fashion to that in London. In Dublin, based on Oscar Faber et al.'s (1998) modelling, somewhere between IR£43 and IR£67 million gross annual revenue would be raised for a morning peak period congestion charging scheme. However, Oscar Faber et al. also estimate that 15-20% of this gross revenue would be needed for operating costs. For the Norwegian cordon toll schemes the approach is different in that the toll is primarily designed to raise revenue for other projects. For instance, in Oslo the gross toll income in 1991 was NOK 600 million, 19% of which is used in operating costs. A further 20% was earmarked for public transport infrastructure and service improvements but the remaining revenue, combined with equal funds from the central government, finances 50 road projects, most of which divert traffic away from the city centre (Ramjerdi and Larsen, 1991). Similarly, in Trondheim the main focus of revenues from the toll will contribute 60% to a package of measures estimated at NOK 2.3bn to increase road capacity (Tretvik, 1993).

The other effect of reducing traffic volumes using road user charging is a reduction of environmental pollution. Cars stuck in traffic can produce three times as much pollution as those driving along a motorway (Institute for European Environmental Policy, 1997). Daniel and Bekka (2000) found that congestion pricing does produce environmental benefits such as cleaner air, less noise and a reduction in energy consumption. The US Environmental Protection Agency (US EPA, 1997) also found that congestion charging schemes could reduce congestion and harmful emissions faster than building infrastructure or changing vehicles. The US Environmental Protection Agency (US EPA, 1998) suggests that the main impacts of congestion charging on vehicle emissions would be expected through effects on vehicle operating speeds. As congestion is reduced and operating speeds become closer to free flow, emission rates per mile of travel would be expected to decrease. For instance, in Edinburgh a reduction in environmental pollutants of between 3.5% and 8.3% has been predicted for



a city centre cordon scheme with a charge of £1stg and £4stg, respectively (Edinburgh City Council, 2001). Santos et al. (2000) also concludes that any toll designed to reduce congestion at peak times would yield the positive by-product of environmental benefits. Newbery (2002) assesses the optimal road user charges from an economic perspective for Britain in 2000, paying specific attention to environmental externalities and finding that in terms of fuel a 'green tax' of 11p/litre on petrol and 15p/litre on diesel would be justified.

Even though the transport effects have been shown to be substantial Mattsson (2003b) has concluded that the locational effects are small. It has been feared that congestion charging might negatively affect the attractiveness of the city centre as a location for shops. However, Mattsson (2003b) has found that in fact the opposite is true with an increase in the number of shops in the city centre, though these results come from a highly simplified stylised model of a symmetric city. Mattsson (2003b) also found that applying a congestion charge on a ring road has a slightly decentralising effect for shop businesses as it is more expensive to use the ring road for shopping purposes. Similarly, decentralisation has been found to result from demand management policies such as parking restrictions, pedestrianisation and bus lanes (Shiftan et al., 2003). The US Environmental Protection Agency (US EPA, 1998) also concludes that congestion charging measures could lead to more dispersed new development or allow better access to currently congested centralised developments, or both. However, Mattsson and Sjolín (2002) have found that service based businesses have a slight tendency to move further out from the city centre. Other studies have also reported small effects; see Eliasson and Mattsson (2001) for a review of the literature.

### ***3.6 Road user charging studies; pre and post implementation***

#### *3.6.1 Introduction*

The cases detailed below cover a wide range of sites, objectives and details of implementation. Hau (1992b), Lewis (1993) and Gomez-Ibanez and Small (1995) have detailed a number of them. They are split into four broad categories:

- (1) Road user pricing of a city centre;
- (2) City centre toll rings designed primarily to raise revenue;
- (3) Road user pricing of a single facility; and,
- (4) Area-wide road user charging.

### *3.6.2 Road user charging of a city centre*

#### **3.6.2.1 Road user charging in the Greater London Area**

London is a city with a population of over 7 million people. Many congestion pricing studies have taken place in the last 30 years or so concerning the Greater London Area. In the 1970s the Greater London Council suggested an area licensing scheme during peak hours. The suggested charge was approximately £1.00stg (1973 prices) to drive in Central London between 8am and 6pm on weekdays. A number of variations to this were also considered with an additional charge for Inner London, which is a larger area surrounding Central London. May (1975) found that congestion would be reduced dramatically with such a scheme and that speeds during the peak period could rise as much as 40%.

Changes to the transport functions of the Greater London Council were taken over by the London Policy Advisory Committee in 1985. In 1988 the Committee proposed a transportation planning strategy with considerably less infrastructural road development than in previous transportation strategies, relying instead on traffic restraint and congestion charging measures. Three concentric cordon rings were proposed, the inner ring surrounding Central London and the outer surrounding Inner London. Screen-lines would also be used to divide up the Central London cordon into 6 zones. A charge of £0.50stg would be levied for crossing either a screen-line or cordon. The charge would apply all day in both directions in Central London and in the direction of peak traffic for the outer two cordons (London Planning Advisory Committee, 1988). May et al.'s (1990) analysis suggests that the proposal would have led to a 15% reduction of inbound traffic to Inner London and a 25% reduction entering Central London.

In 1994 the UK Department of Transport commissioned a comprehensive three year study of congestion charging for the Greater London Area conducted by MVA Consultancy (MVA, 1995). A number of charging schemes were investigated as part of the London Congestion Charging Research Programme (MVA, 1995) but the simplest one was a single cordon charge for Central London with 130 charge points. It was found that the application of an inbound charge of between £2stg and £8stg from 7am to 7pm would provide between an 8% and 22% reduction in the total vehicle kilometres travelled in the city. The estimated net benefit of the scheme after taking into account time savings, implementation costs, accident cost savings, etc. ranged from approximately £35stg million to £60stg million (Small and Gomez-Ibanez, 1998). The scheme also offered high reliability, no disruption to traffic flow and protection of privacy for users, making it a more acceptable scheme to the general public. However, at the conclusion of the study the Minister of Transport declared that congestion charging would not be undertaken in London until at least the end of the 1990s.

In 1999 the Greater London Authority was created under new legislation and a strategic proposal for transportation in London, 'Transport Strategy for London', was produced by the mayor, which included congestion charging (Dix, 2002; GLA, 2001). The suggested charging scheme was based on the 'Road Charging Options for London' (ROCOL) study produced by the 'Review of Charging Options for London Working Group' and commissioned by the Government Office of London (GOL, 2000). The scheme is an inner city cordon with a £5stg charge operating between 7am and 7pm on weekdays within Central London, see Section 3.4 and Section 3.5 for acceptability and effectiveness.

The Central London congestion charging cordon scheme has now been in place for a period of six months, operating from 7am until 6.30pm with a £5stg charge five days a week. A three month report has been produced (Transport for London, 2003) that describes the effectiveness of a city centre congestion charging cordon scheme. Overall, a reduction in congestion of 32% has been achieved. In addition, the scheme has resulted in a 16% reduction of traffic levels inside the cordoned Central London

area. When compared with the 10-15% estimation in the Rocol report (GOL, 2000) these results indicate that the scheme is more effective than modellers first estimated.

### **3.6.2.2 Road user charging in Singapore**

Singapore was the first to implement a working congestion charging scheme, commonly known as Singapore's Area Licence Scheme (ALS), which was inaugurated in 1975 and still operates today. Its initial system was very simple with the priced area (over 5km squared) defined within a single city centre cordon with a charge of S\$3.12 (Holland and Watson, 1978; Chin, 2002). Paper windshield stickers were used and enforcement was carried out through visual inspection by traffic officers.

The ALS was designed to restrict car ownership and the demand for road space on the crowded island city-state. Congestion charging was chosen over other means of demand management because of a lack of space for tolling stations in the city centre and because it was thought that higher parking charges would have little effect against heavy through traffic and a high percentage of chauffeur driven cars.

When implemented in 1975 the charge was applied to inbound traffic crossing the cordon in the morning peak period, though this charge period from 7.30 to 9.30 was extended by 45 minutes because so many people were delaying their trips until after the charged period (Willoughby, 2001). The charging period was extended further in 1989 to include the afternoon peak, and further again in 1994 to include the time between the two peak periods. Charges of S\$3 and S\$2 were used respectively (Chin, 2002).

Traffic demand was effected in the desired way. For instance, commuters to jobs in the restricted zone travelling by car with less than four passengers dropped from 48% to 27% while the combined modal shares of car-pool and bus rose from 41% to 62%, see Holland and Watson (1978). Traffic volumes also dropped by 45% and the average speed increased from 18 to 35km/hr.

Some latent demand was initially released with the frequency of uncharged trucks entering the pricing area during the peak period increasing by 124% (Watson and Holland, 1978). However, once the charge was extended to all vehicles, except public

transport and emergency, in 1989 their use declined immediately. These responses demonstrate that very specific reactions can be garnered through pricing initiatives, while any loopholes are likely to be heavily exploited (Small and Gomez-Ibanez, 1998).

Finally, in 1998 the current paper permit system was replaced by an electronic smart card system. The ALS was replaced by an Electronic Road Pricing (ERP) scheme and the objectives of the scheme were changed to improve travel speeds on the road network, which was predominantly to be achieved by forcing a transfer from private cars to other modes (Menon and Chin, 1998; Menon, 2000; Goh, 2002). Each vehicle entering the restricted zone, which has 33 gantries for control and enforcement placed at the boundary of the zone, during the charge period between 7am and 7pm on weekdays and 7am and 2pm on Saturday, must have an “In-vehicle Unit” (IU) with a smart card in credit inserted. The charge is then automatically deducted as the vehicle passes under a gantry. The charge is varied between S\$0.50 and S\$2.50 to maintain a traffic speed of 45-65km/hr on expressways and 20-30km/hr on arterial routes (Chin, 2002; Goh, 2002). This has resulted in a reduction in traffic volumes within the restricted zone of 20% to 24%. The ERP system is more flexible, easier to use and technically more reliable than the old ALS system but has a very high investment cost, which is deemed worth it in the case of Singapore where congestion control is a priority (Foo, 1997, 2000). However, Foo (2000) also concludes that the manual ALS system has laid the foundations for the more technologically advanced ERP system to be a success.

### **3.6.2.3 Hong Kong’s Electronic Road Pricing Trial**

While Singapore was the first to implement a congestion charging scheme Hong Kong, a slightly larger city with a population of approximately 4 million, was the pioneer in fully-automated charging technology using an ERP scheme with multiple cordons (Harrison et al., 1986). However, while the ERP system was subject to extensive field trials and transportation modelling studies for prediction and evaluation, it was withdrawn due to strong public opposition.

These field trials demonstrated that ERP systems could operate with a high degree of accuracy. Subsystems for automatic charging, billing, and enforcement through closed-

circuit television all performed very well (Dawson and Catling, 1986). Transportation modelling studies of alternative pricing structures and charging locations were undertaken by MVA Consultancy (Harrison, 1986).

Three different schemes were studied of varying size using a cellular zoning system and time varying charges. The average monthly payment for the three schemes was approximately HK\$120, HK\$140 and HK\$160, respectively (Transpotech, 1985). A reduction in demand of between 20 and 24% was predicted during the peak periods with an overall daily reduction of 9% to 13% (Harrison, 1986). Taking the most complex scheme as the benchmark, the simplest scheme (A) appeared to approximate marginal cost pricing well. Scheme B had more of an effect on congestion by charging more in the peak direction, which appears to be more important than refining the geographical layout (as in Scheme C) (Small and Gomez-Ibanez, 1998).

The Hong Kong ERP scheme did not come to fruition due to a lack of public and political acceptance. There were a number of reasons for this (Small and Gomez-Ibanez, 1998):

- (1) The trials took place during the transfer from British colonial government to local officials;
- (2) A weak economy had lowered car ownership in the early 1980s;
- (3) Objections due to the potential invasion of privacy through electronic equipment (Hau, 1990, 1992b; Gomez-Ibanez and Small, 1995); and,
- (4) The public did not perceive that the revenues gained through such a scheme would benefit them.

In March 1997 a feasibility study on an ERP System (ERPS) was undertaken by the Hong Kong SAR Government, assessing two technology options, a Dedicated Short Range Communications (DSRC) System, and a Vehicle Positioning System (VPS). Their analysis found that peak hour charges should be applied from 8am to 9am and from 5.30pm to 7pm with a slightly lower rate during inter-peak times. The charges

used ranged from HK\$8 to HK\$35 depending on the time of day and the traffic demand growth scenarios (Hong Kong SAR Government, 2001). The field trials found that either the DSRC System or VPS could be adapted for possible implementation in Hong Kong. An estimated 40% of car trips in the morning peak may be diverted to public transport with 10% changing travel times so as to avoid the charge (Hong Kong SAR Government, 2001). However, a stated preference survey involving a HK\$20 charge led to a 13-15% mode switch to public transport and 24-27% of car users changing their time of travel (Transport Bureau, 1998). The feasibility study concluded that a congestion charging scheme is not warranted on traffic management grounds before 2006 if the growth of the private vehicle fleet is no more than 3% per year (Hong Kong SAR Government, 2001).

#### **3.6.2.4 Congestion specific pricing for Cambridge, England**

In Cambridge, theoretically optimal congestion charging was attempted, which closely approximated the theoretical ideal of road user pricing by varying the charge in real time to reflect the level of congestion encountered by the individual vehicle while travelling within the priced area (Sharpe, 1993). The rationale for the idea was that the congestion experienced by drivers is probably closely related to the externalities it imposes on other vehicles. Field trials were undertaken but the scheme was abandoned after a new, less supportive, city council came to power in 1993.

A metering system was to be used where technology was tested as part of the ADEPT project within the European Union's DRIVE-II program (Clark et al., 1994). In 1994, after the implementation of the scheme was abandoned, surveys found that only one third of respondents described the road user charging concept as 'acceptable' (Ison, 1996). May et al. (1994) have found that the use of congestion specific charges do significantly increase the benefits over and above those gained through a cordon type system. However, Sharpe rejected the congestion specific idea in early 1993 because of concerns about the potential for public disapproval where charges appeared to be unpredictable. His concerns proved to be well founded with the end of the proposed scheme later in 1993 as a result of a change in local shire government. Ison (1998)

describes several reasons why congestion metering failed to advance beyond the field trial, concluding that the timing of the project and the level of sophistication were important contributing factors to its downfall. In a further paper Ison (2000) concludes that for Councillors (local government officials) public transport improvement, lower fares, heavier restriction of vehicles in central areas and land use and transport planning strategies are viewed as more effective and socially acceptable.

Cambridge has, however, demonstrated the technical feasibility of more complex forms of road user charging. Unfortunately, it has also demonstrated the need to develop support for road pricing schemes in parallel with any design proposals if the scheme is to prove acceptable, both socially and politically.

### 3.6.3 *City centre toll rings; the Scandinavian experience*

A number of Scandinavian cities have developed road pricing as a tool for raising revenue to finance road infrastructure. They are described not as congestion management schemes but primarily as revenue generators. As such, the toll schemes that have been implemented use a low charge and do not vary much by time of day. It is also noted that cordon locations were chosen to be both politically acceptable and to achieve the correct balance between the contributions of both city and suburban residents, while altering trip behaviour as little as possible. It has been noted that public confidence has increased with practical experience as long as the schemes have remained closely tied to well known and widely shared objectives (Small and Gomez-Ibanez, 1998).

#### **3.6.3.1 Norway's urban toll rings**

Toll rings are an extension of the single facility tolling strategy used to raise money for road infrastructure. Instead of applying a toll to finance a single facility, such as a town or city by-pass, the toll is applied over entire urban road networks. Each toll ring is part of a regional financial package of major road development. As a result Norway's tolls are low, ranging from NOK5 in Bergen to NOK10 and NOK 11 (1992 prices) in Trondheim and Oslo, respectively (Small and Gomez-Ibanez, 1998).



Bergen has a population of approximately 300,000. In 1986 it opened a manual system operating from 6am to 10am on weekdays (Larsen, 1988). Oslo, the nation's capital with a population of approximately 700,000, followed four years later in 1990 with a system of 19 toll stations operating at all times. In 1991 Trondheim introduced a more complex system operating for 11 hours per day, 6am to 5pm, on weekdays with a discount after 10am.

In Trondheim electronic subscribers account for approximately 95% of all tolled crossings and benefit from a discount after 10am, and, from ceilings on their charge liabilities in any given hour or month. These characteristics could, according to Small and Gomez-Ibanez (1998), enable the system to approximate congestion charging. Literature published by the Public Roads Administration even makes reference to this fact.

New charges have been put forward for Trondheim with three new screen-lines planned for inside the large outer ring. This produces three zones and, combined with an increase in the toll of NOK1 for crossing the outer ring and an extension of one hour to the operating time (to 18.00), is expected to lead to an increase of 50% in toll revenues. It is noted that the charge for crossing internal screen-lines is considerably less than for crossing the main outer ring.

The effect of these pricing schemes is as expected with inbound vehicle crossings reduced by no more than 6% to 7% in Bergen, 8% in Oslo and 10% in Trondheim (Larsen, 1995). However, in Trondheim 40% of the general public indicated some kind of change in their behaviour, be that mode change, time change, trip frequency, etc. (Meland and Polak, 1993). Odeck and Brathen (2002) assess the future of toll pricing in Norway concluding that the public needs to be more involved in the planning process prior to implementation, the tolls offer great potential to counteract the environmental damage caused by traffic, and, that there are potential economic benefits to introducing Public Private Partnerships (PPPs) for the transport sector.

### 3.6.3.2 Road user charging in Stockholm

Stockholm is twice the size of Oslo. It is the capital city of Sweden, which has a strong environmental movement concerned with problems associated with traffic in inner cities (Small and Gomez-Ibanez, 1998). In 1991 negotiators from the three political parties agreed to invest approximately \$6.9bn over 15 years in urban transportation improvements to be financed primarily by road pricing. The program, known as the Dennis Package, devotes more than half of the funds to road improvements and the balance to public transportation (Social Democratic Party, Moderate Party and Liberal Party, 1991).

The project involved the completion of an inner ring road, the construction of a tolled north-south bypass to the west of the city centre and an inner city toll ring designed in part to reduce traffic volumes, noise and air pollution and to raise revenue. Johansson and Mattsson (1995) found that these elements would combine well to reduce vehicle travel in inner Stockholm, mitigating the effects of additional traffic caused by construction of the new roads. Pre-purchased licences were to be used for the toll operating on weekdays with £30 monthly smart cards or £2.50 daily passes resulting in a predicted reduction of traffic of up to 10%. Construction was to begin in 1997 but failed because of strong public opposition (Ahlstrand, 2001).

Recent elections in 2003 have resulted in another change of government and a change of opinion with reference to road user pricing has followed (Mattsson, 2003a). Consequently, a full-scale test of congestion pricing is to be carried out before the end of the present term of office, i.e., 2006. This test is to start by the end of 2004 and a referendum will be held at the time of general elections in 2006 to determine if it is to be made permanent or not (Mattsson, 2003a).

A simple electronic zone-based system with 2 zones is proposed with the inner city divided between North and South. The charge period will be on work days between 7am and 6.30pm and the charge level is suggested to be 20 SEK for crossing the cordon boundary in any direction during peak hours and 10 SEK during off-peak. The fee for passing between North and South is suggested to be 10 SEK during peak periods. The

collected revenues are to be used as investment in the public transport system in the Stockholm region.

Mattsson (2003a) reports on three studies of congestion pricing in Stockholm as part of an ex ante study of transport to provide a knowledge background to the system being considered at present. For the zone-based system, traffic volumes in the inner city are predicted to drop by 30% for charged hours and a charge equivalent of 3 SEK/km. For the distance-based system, traffic volumes in the inner city are predicted to drop by 35% and 19% at charge levels of 4 and 2 SEK/km, respectively. Thirdly, in the case of optimal congestion pricing, the reduction of traffic levels is 25% at an average charge level of 2 SEK/km. Locational effects were also studied and were found to be very limited (Mattsson, 2003a), see Eliasson and Mattsson (2001). In the opinion of Mattsson (2003a), the use of revenues, including guarantees that the money will be channelled back into the Stockholm region, is the key issue in gaining political and public acceptance for congestion pricing in Stockholm.

#### *3.6.4 Area wide road user pricing; The Randstad Region, The Netherlands*

The Randstad region of The Netherlands is an urban agglomeration covering 2000 square miles with a population of over 6 million. Included in the Region are the cities of Amsterdam, Rotterdam, Utrecht and The Hague. In the 1980s the government proposed a road pricing development for the region. The proposal involved 140 charging points with time-varying tolls across a multi-cordon system. A 17% reduction in congestion was modelled using a charge of 5 guilders. However, publicity had focused on the technical aspects of the scheme rather than on the benefits to the general public so in 1990, the government, unable to raise the levels of approval for the scheme, reduced the scope of the proposal calling for more conventional road tolls instead (Grieco and Jones, 1994; Emmerink et al., 1995).

The Ministry of Transport and Public Works revived the idea of congestion charging in 1992. The new proposal called for congestion charging along the main arterial road system during the morning peak period from 6am to 10am. The proposal was set back in 1994 with the election of a new government but in 1995, the new Minister for

Transport suggested implementation starting in the year 2001 for the sole purpose of congestion management. The Minister also mentioned a possible 50% reduction of congestion during the morning peak period on the main road network.

In June 1995, the Minister surveyed peak period road users to assess public opinion on the proposal (Verhoef, 1996). Fifty percent of those surveyed thought that road user charging was a bad idea, 25% thought it was a good idea while 85% said their opinion would depend on the allocation of revenues from the scheme. Of these allocations road investments and vehicle-related tax reductions were at the top of the list, closely followed by public transport investment. However, privacy was also a public concern in the case of an ERP scheme (Phang and Toh, 1997).

In 1998, Mu-consult modelled a kilometre tax charge system to replace private car registration tax (BPM) and annual vehicle tax for ownership (MRB) (Mu-consult, 1998). A number of scenarios were modelled using distance based charging of between 3.1 and 9.1 Euro per kilometre resulting in a reduction of between 6% and 20% of private car volume (Boot et al., 1999). The effects of a large increase in fuel taxes were also assessed (50% increase) resulting in only a 6% reduction in traffic volumes and the conclusion that increasing fuel taxes is far less effective than distance-based charging systems (Boot et al., 1999).

In 2001 a new charging scheme was proposed by the Dutch Ministry of Transport, Public Work and Work Management in its strategic proposal for transportation up to 2020 entitled, 'The National Traffic and Transport Plan' (Sociaal-Economische Raad, 2001). The road pricing project began with the publication of the 'Mobimiles' report by the Pieper committee, which recommended an implementation strategy and framework for a variable road pricing system. The charge could vary according to location, time of day, and, environmental parameters determined by the type of vehicle (EUROPRICE, 2002). This would be made possible by using an in-board unit called the 'mobimetre'. However, the project failed to go ahead once again because a new unsympathetic government was installed in July 2002. This government's strategic approach is to

focus on broadening the road infrastructure and improving the public transport network (EUROPRICE Newsletter, 2002).

### *3.6.5 Single facility road user pricing: California's private toll lanes*

The first example of congestion charging, or value pricing as it is known in the United States, is located in Southern California along the median strip of the existing Riverside Freeway, otherwise known as State Route 91. The scheme used has come to be known as 'HOV buy-in' or high occupancy toll (HOT) lanes (Fielding and Klein, 1993). The unused capacity in HOV lanes is auctioned off to low occupancy vehicles.

Further to this idea, the operator is considering toll rates that would vary in response to real-time measurements of congestion levels. Information about the level of charge and the delays on the free lanes running in parallel would be given on existing message signs. The users would then know the price in advance of their decision to take the HOT lane as well as the congestion situation on the free lanes. This represents a further development from the real-time scheme proposed in Cambridge where users would not be aware of the changing charges.

The scheme received a 60-70% approval for the use of toll finance and a 50-60% approval, up from 40-50%, before introduction of time varying tolls (Small et al., 1997). Delays on the free lanes dropped dramatically with the introduction of the scheme using tolls of \$0.25 to \$2.50 varying by time of day. Additionally, it was found by Sullivan (1996) that latent demand was not released onto the corridor either from parallel routes or from newly generated traffic. Small and Gomez-Ibanez (1998) believe this reflects the mountainous terrain and the resultant lack of alternative routes.

## **3.7 Price elasticities of user demand**

### *3.7.1 Introduction*

The user response to road user charges, i.e., the price elasticity of demand for car travel, is an uncertainty that needs to be recognised as there are many factors that affect it. Important factors influencing the calculation and interpretation of price elasticities

include: the size of the price change; the type of pricing mechanism; the type of trip; the type of traveller; the price of related goods and services; and whether the elasticity accounts for short term or more long term demand responses. Hence, care should be taken when applying price elasticities that they are based on a similar context to that in which they are being applied. In the case of a road user charging scheme analysed in the long term, say 10 or 20 year predictions, the elasticity associated with the scheme changes over time. Initially people's reaction is slow to the introduction of such a scheme so the user responsiveness is low, hence a low elasticity. This elasticity can increase from anywhere between 50% and 200% during the change from short to long term forecasting (Cole, 1998).

### 3.7.2 *Review of pricing elasticities*

Although there are a limited number of road pricing studies a large body of literature does exist on the estimation of demand elasticities, and a number of literature reviews have been published which summarise this literature very well. These reviews include Oum et al. (1992), Goodwin (1992) and Glaister and Graham (2000), and look at evidence about increases of fuel price, parking charges and tolls. These can, to a certain extent, be considered transferable to express effects of road user charging. It is suggested in Oscar Faber et al. (1998) that a peak demand elasticity value of between  $-0.1$  for low charges (IR£1), rising to  $-0.5$  for higher charges (IR£3) be used for the Greater Dublin Area (GDA). For other examples of road user pricing elasticity estimates, varying from  $-0.1$  to  $-1.2$ , see Cole (1998), Oldfield (1981) and Halcrow Fox (1993).

Elasticities of  $-0.18$  and  $-0.36$  are used by Santos (2000b) in an estimation of whether the benefits of introducing road pricing would exceed the costs of introducing it by enough to justify its introduction. The elasticities were justified by the fact that they are roughly in line with fuel price elasticities in Oum et al. (1992) and Goodwin (1992). Also, these elasticities represent short run elasticities of demand concerning fuel costs and time costs. As such, they are probably greater than fuel price elasticities (Santos, 2000b).

Menon (2000) has calculated the short-term demand elasticities, for the 8am to 9am peak period for cars in the restricted area, to be between  $-0.12$  and  $-0.35$  for Singapore's ERP scheme in 1998. This compares to the previous ALS in 1976 where the elasticity figures stabilised at  $-0.15$ . However, Luk (1999, cited in Jaensirisak, 2002) found that for the area licensing scheme in Singapore toll elasticities varied between  $-0.19$  and  $-0.58$  with an average of  $-0.34$ , which is twice as effective as increases in petrol prices. Further to this Luk found that in relation to mode shifts road pricing was as effective as increases in petrol price. However, Luk and Hepburn (1993) found that in the long term changes in petrol price had more effect in reducing petrol consumption, an elasticity of  $-0.55$ , than in reducing travel demand, with an elasticity of  $-0.26$ .

In the case of Stockholm's proposed road pricing scheme, Mattsson (2003a) calculates arc elasticity of vehicle distance travelled with respect to total vehicle operating costs, leading to an elasticity for a distance-based road pricing scheme of  $-0.23$ . Here arc elasticity refers to the distance between two points on a curve produced by graphing the vehicle distance travelled with respect to total vehicle operating costs. It is noted by Mattsson (2003a) that the value of  $-0.23$  is well within intervals described by Goodwin (1992) and does not represent an extravagant assumption about the trip makers' cost sensitivity. This value is also similar to experiences in Norway where elasticities of  $-0.22$  and  $-0.45$  were found for Oslo and the Alesund toll station, respectively (Jones and Hervik, 1992).

A study has been carried out by Arentze et al. (2003) using 'Albatros', a multi agent system developed by the Dutch Ministry of Transport that predicts which activities are conducted where, when, for how long, by whom, and the transport mode involved. For congested roads a price elasticity of travel demand in the range  $-0.35$  to  $-0.39$  was found, though it is noted that stated models tend to overestimate choice flexibility. Oum et al. (1992) summarise price elasticities found in studies from various countries, and are mainly concerned with fare charges. The elasticities found lie in the range of  $-0.09$  to  $-0.52$ , though the context and response alternatives differ in several ways to those used by Arentze et al. (2003).

Halcrow Fox and Associates (1993) performed a study of road pricing in London using stated preference surveys in 1992. The study assessed the effects of distance based charges on private cars within Inner London. It was reported that the elasticities for radial work journeys in the morning peak, being charged at a rate of 5, 15 and 35p/mile, were  $-0.11$ ,  $-0.38$  and  $-0.52$ , respectively. Similarly, for orbital work journeys the elasticities were  $-0.06$ ,  $-0.12$  and  $-0.41$ , respectively. The difference between radial and orbital elasticities was explained as the result of better public transport services on radial routes affording better alternatives to car use.

In 1994 the SACTRA (Standing Advisor Committee on Trunk Road Assessment) Report on Induced Traffic was produced (UK DETR, 1998b). Following this MVA conducted research on Induced Elasticities and Methods (IEM) derived from the SACTRA Report, as part of a study for the UK DETR relating to possible responses to increases in congestion. These results were reported in 'Effects of Congestion Final Report' (MVA, 1995). Further to this, the 'Phase 4 Report: Re-analysis of Effects of Congestion Data' was produced by MVA (UK DETR, 2000), which focuses on private responses to increased traffic congestion using the stated preference survey results from 1995 (MVA, 1995).

A trip mileage elasticity of  $-0.14$  is given for work related trips in London, which represents both short and long-term effects. Short and long-term time elasticities, allowing time switching to avoid the charge, of  $-0.27$  and  $-0.33$ , respectively, are also reported for work trips in the morning peak. These elasticities are supported by the ROCOL study (GOL, 2000) where elasticities of  $-0.09$  and  $-0.27$  are reported for high and low income groups, respectively, where stated preference responses to a £10stg Inner Ring Road charge were used. ROCOL (GOL, 2000) have also used elasticities, found using the AREAL model (a strategic model developed by the Government Office for London to evaluate the impacts of alternative charging options in the Greater London area), which relate to work trips with respect to £2.50stg and £10stg charges, of  $-0.17$  and  $-0.27$ , respectively. Elasticities of  $-0.20$  and  $-0.31$  were also found for peak period trips including those related to education, work, employer's business and other.



The UK Department for Transport has used a 'car traffic with respect to fuel cost demand' elasticity of  $-0.3$  for the demand models calibration year of 1998 (Transport for London, 2003). An elasticity of  $-0.5$  was used by Stoneman (1998) for network modelling carried out as a part of the "Transport Demand Management in Historic Cities" research project. The National Economic Research Associates (NERA, 2002) suggest the use of elasticities of at least around  $-0.1$  to  $-0.3$  in the short run for road user charging schemes based on distances travelled and levels of congestion. There is no evidence that the effects of petrol price and road user charging are the same yet an assumption is made that they are unlikely to be very different. This, according to NERA (2002), is based on the fact that while some travellers may avoid charges by switching to periods with less or no charge and, hence, not reduce their use of the car, others may drive more efficiently in order to mitigate the impact of a fuel price increase.

Glaister and Graham (2000) report short run elasticities, in particular, traffic levels with respect to petrol price of  $-0.15$  and a price elasticity of demand for petrol consumption of  $-0.3$ . These are in line with Goodwin (1992) who found the price elasticity of demand for petrol consumption to be  $-0.28$  and traffic levels with respect to petrol price to be  $-0.16$  and  $-0.33$  in the short and long-term, respectively. Goodwin also reported an overall elasticity of  $-0.46$ . Jong et al. (1999) suggests that fuel price elasticities could be used for evaluating distance-based charges in particular, and that area-based or cordon-based road pricing could be evaluated using parking charge elasticities because the charge is more or less fixed per trip. The average parking charge elasticity in Western European countries was found to be  $-0.10$ . However, Jaensirisak (2002) suggests by way of an explanation, that because many commuters have free parking, parking charge elasticities could be lower than area-based pricing elasticities. Jong et al. (1999) continue on to report a fixed price elasticity for car commuting trips of  $-0.20$  in the short run for Western European countries, which is reasonably in line with other studies.

### 3.8 Conclusions

#### 3.8.1 Effectiveness of road user charging

Through the case studies and literature discussed above, using either preliminary studies or reports from implemented schemes, it has been demonstrated that road user charging can substantially affect user behaviour and hence, traffic behaviour based on modelling studies. Road user charging can be used to raise revenue, reduce environmental pollution and reduce congestion. With the reduction of car trips during peak periods a transfer of trips to other modes can be expected as the primary response to a road user charging scheme. Hence, to be equitable revenues collected need to be re-allocated to benefit those charged, which can be done by reducing motor related taxes, improving public transport services and infrastructure, etc.

Locational impacts appear to be very small leading to the conclusion that geographical layout plays only a minor roll in the effectiveness and impact of road user charging schemes. However, from a business perspective congestion charging of a city centre tends to increase the number of shops within the city centre, as opposed to decreasing the number as was initially feared.

Overall, international experience appears to be very positive. However, pitfalls are especially prevalent when dealing with political and social acceptability of congestion charging schemes, i.e., the experiences in Stockholm, Hong Kong and Cambridge. With respect to the implementation of congestion charging schemes a city centre cordon scheme with a fixed charge appears to be favourable, especially in relation to social and political acceptance. Once the public has been given the opportunity to get used to such a scheme then it is far more likely that they will accept a more complex upgrade like that experienced in Singapore.

#### 3.8.2 Political and social acceptability of road user pricing

In order to implement a successful charging scheme it is, therefore, necessary to design the scheme not just to be effective in achieving its objective but to be acceptable to the general public as well. Further to this, Jaensirisak (2002) has found that acceptable road

user charging schemes can be designed to be both acceptable and effective by limiting the area of charge to within the city centre and having a fixed charge per day. It was also noted that support would increase significantly if there was substantial improvement to the environment. Goh (2002) suggests that even though some road user pricing systems may result in a decline in overall welfare for some users (Phang and Toh, 1997), for instance in Singapore, this can be mitigated by improving alternative transport modes that can operate independent of road conditions. Therefore, any road user charging scheme must be implemented as part of an integrated strategy that combines transport demand management with infrastructure and service improvement.

### *3.8.3 Road user charging studies; pre and post implementation*

The design and implementation of congestion charging schemes is very important for public acceptability, particularly with respect to charge levels, enforcement, privacy and the degree of public understanding. Political acceptability is difficult to attain because most users tend to lose out. To counteract this feeling of loss it is necessary to make users perceive the benefits of charge revenues by using them to finance transportation improvements, or to lower vehicle associated taxes. Another approach for overcoming peoples suspicion of change is to apply a strategy of incremental development (Small and Gomez-Ibanez, 1998). For example, the Norwegian toll rings began as a means of raising revenue for financing regional transportation infrastructure but have begun to incorporate traffic management as a subsidiary goal.

At a more advanced level it is evident that introduction of a simple congestion charging cordon scheme in the first instance is preferential if the scheme is to be a success (GOL, 2000; Oscar Faber et al., 1998). More complex systems are more likely to be successful after a simple system has been seen to work by the general public. In Singapore a very simple congestion charging cordon scheme was introduced first and has since been developed into a more complex system with multiple cordons and screen-lines over a wider area.

Charging levels have varied greatly across the case studies described from \$0.25 to £8stg. However, according to Santos (2000a), errors in estimating the effective charge

will not hugely affect the efficiency of the system. Therefore, based on implemented examples of congestion charging, i.e., Singapore and London, a charge in the range of 3 to 8 Euro would be appropriate.

#### *3.8.4 Price elasticities of user demand*

At this stage it is not possible to assign specific elasticities to represent user response to road user charging scenarios in the GDA. Given the need to make an assumption on elasticity values, they should be restricted to within a close range that is reasonably justified. Taking into account the literature and studies described in Section 3.7, such as Goodwin (1992), Oum et al. (1992), Jong et al. (1999), Glaister and Graham (2000), the UK DETR (2000, 2001), as well as practical experience from, for example London, a sensible range of values might be considered to be -0.2 to -0.7 for an inner city fixed price congestion charging scheme.

## *Chapter 4*

# *Improving the Performance of the Dublin Transportation Network Model using High Performance Computing Techniques*

#### **4.1 Introduction**

The transportation environment in the GDA has been described in Chapter 2. This chapter demonstrates that there is a significant transportation deficit in the GDA. The DTO's strategic plan for the period 2000 – 2016, 'A Platform for Change' explicitly recognises this fact (DTO, 2000). The DTO (2000) have also recognised that infrastructural development alone will not provide a sustainable transportation environment in the GDA and that some form of demand management will have to make up an integral part of any successful transportation strategy for the future of the GDA. Success will be quantified by the measure of sustainability and accessibility, through a reduction in travel times and congestion over the private car and HGV transportation networks from the congestion levels experienced in 1991, when the average speed in the morning peak hour was 22kph; see Chapter 2 Section 2.6.2 for more detail. Considering that at present the trip levels in the GDA are already at or exceeding the levels predicted for 2016 in 'A Platform for Change' it has become imperative that the potential of demand management, its implementation and the effects of any potential implementation for the GDA are investigated.

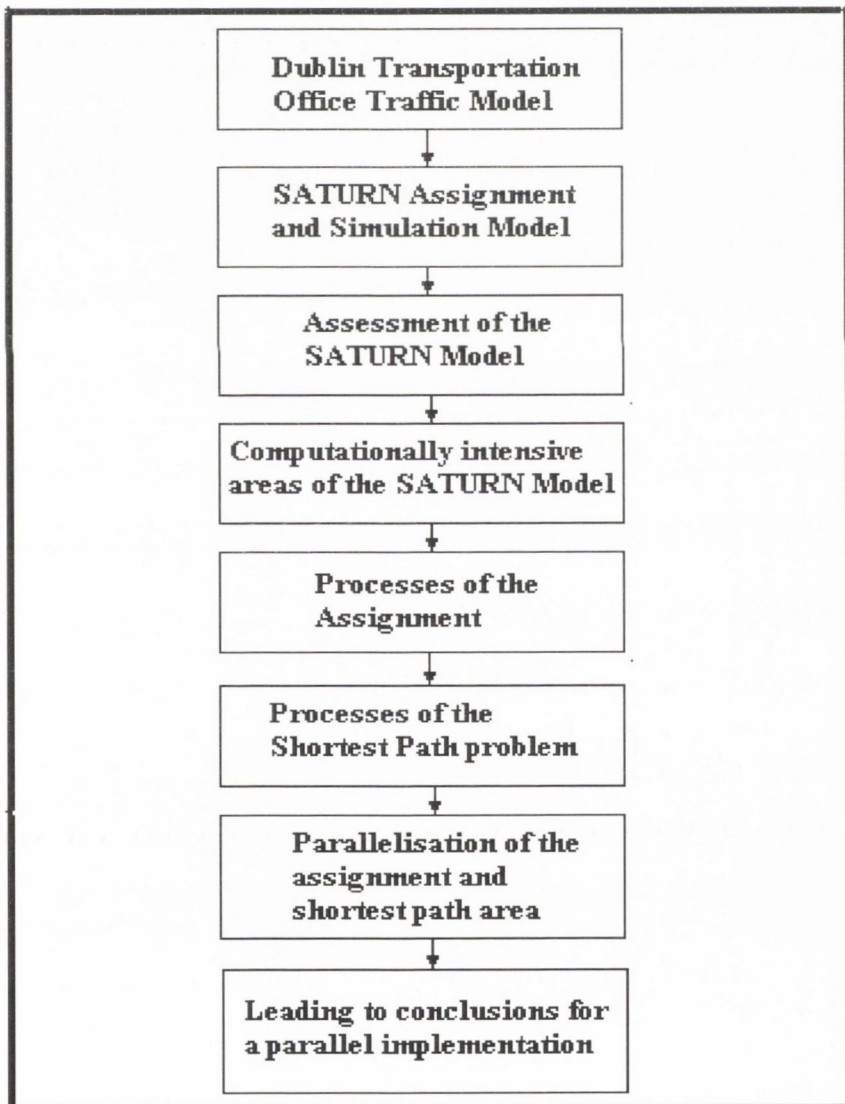
As Chapter 3 describes, there are a number of instruments that can be applied to manage the demand for transport in urban areas. However, it is found that all of these instruments, except for road user charging, fail to target the external costs of urban motoring adequately (Oscar Faber et al., 1998). Therefore, Chapter 3 goes on to give a review of road user charging as a means of demand management and describes any implications that this review has for this research. It is concluded that an inner city cordon scheme with a fixed rate of charge would have the greatest potential for succeeding to reduce the demand for travel by the required levels in order to achieve the objectives laid out in 'A Platform for Change' for the GDA. It is also found that an investigation into the effectiveness of any such application, or variation there of, is required before it would realistically be considered. This type of investigation would take a considerable number of model runs given the area to be covered and the various combinations to be tried. Undertaking this investigation was not practical at the time due to the number of model runs required and the length of time taken to execute a

model run. It was, therefore, concluded that an improvement in the running time of the DTO's Traffic Model was necessary to conduct research on possible road user charging initiatives within the GDA.

This chapter will research how to improve the performance of the DTO Traffic Model in the Greater Dublin Area with the aid of high performance computing techniques and hardware. The objectives of this research for the DTO's modelling process can be described as follows:

- To investigate the possibility of reformatting the code of the SATURN (Simulation and Assignment of Traffic to Urban Road Networks) model, see Section 4.3.2, in order to incorporate parallel processing and, ultimately, reduce the processing time for each model run. This would represent a very useful enhancement for the DTO Traffic Model and the other authorities currently using the model. It would also have implications for other users of the SATURN Suite in the U.K. and Europe;
- To make use of the IBM 9076 Scalable POWERParallel SP2 supercomputer, shared between Trinity College Dublin and Queens University Belfast, and the computing power available by producing new improvements to the SATURN model using parallel processing methods. This may be achieved by concentrating on the algorithmic loops both between and within the assignment and simulation procedures of the SATURN model. The long run times of the SATURN model are due to the numerical computation involved in the assignment and simulation of traffic conditions, such as those defined by the transportation network for the GDA. Using the far superior computational capacity of the IBM SP2, it is hoped that the long model run times may be cut by a significant amount allowing the SATURN model, in conjunction with the DTO Traffic Model, to be used in a wider range of problems than it is presently feasible to undertake; and,
- To use the new model to conduct research on possible road user charging initiatives within the GDA, which it is not practical to undertake at present due

to the long run times of the model. This type of investigation would take a considerable number of model runs given the area to be covered and the various combinations to be tried. With the improved performance of the parallelised model this could then be achieved within a reasonable period of time due to the shorter run times of the new parallel model. This, according to Hribar et al. (2001), is the primary goal of using parallel machines for traffic equilibrium applications.



**Figure 4.1:** Outline of the areas assessed in Chapter 4 and the progression through to a set of conclusions for parallel implementation of the SATURN model.



Figure 4.1 gives an outline of the areas assessed within Chapter 4. A summary of the more important and practical aspects of the DTO Traffic Model is described in Section 4.2. Section 4.3 briefly describes the SATURN traffic assignment and simulation model. Section 4.4 describes the initial analyses of the SATURN model, assessing if potential performance improvement can be achieved through parallel processing, and if so, where it might be achieved. Section 4.4 concludes that the computationally intensive area of traffic assignment has the most potential for performance enhancement. As a result, Section 4.5 describes the technical aspects of traffic assignment within the SATURN model leading to a description of the shortest paths problem and tree building in Section 4.6. After thoroughly investigating the areas within the SATURN model where parallel processing could give relevant performance enhancements, a literature review of parallel processing in those areas is undertaken in Section 4.7. Finally, Section 4.9 summarises the more important aspects and conclusions of the chapter.

## ***4.2 The Dublin Transportation Office Traffic Model***

### *4.2.1 Introduction*

The Dublin Transportation Network Model (DTNM), or the Dublin Transportation Office (DTO) Traffic Model as it is better known now, was first set up as part of the DTI strategy to assess strategic transportation proposals in the DTI Study Area (DTI, 1995). This modelling process has been described in Chapter 2 Section 2.4.4. The DTNM is described as a zone-based local model where strategic zones are addressed individually. The DTNM made use of disaggregate data such as car ownership and car availability and the trip making characteristics of the population, all of which were designed to be easily updateable and enhanced in the future.

The model has since been updated by the DTO in 1996, 1997 and again in September of 2002. The DTO Traffic Model uses one of the largest representations of the SATURN model in Europe assessing medium to long-term prospective strategic transportation planning proposals in the GDA.

The section that follows will summarise the more important aspects of the DTO Traffic Model. Section 4.2.2 looks at zoning within the GDA. Section 4.2.3 describes transportation networks; specifically, how they are defined in terms of zones, links and nodes. Section 4.2.4 describes the classic 4-stage model structure looking at trip generation, distribution, modal split and assignment from a practical standpoint. Finally, Section 4.2.5 describes the overall model structure while Section 4.2.6 concludes briefly.

#### 4.2.2 Zoning areas of interest

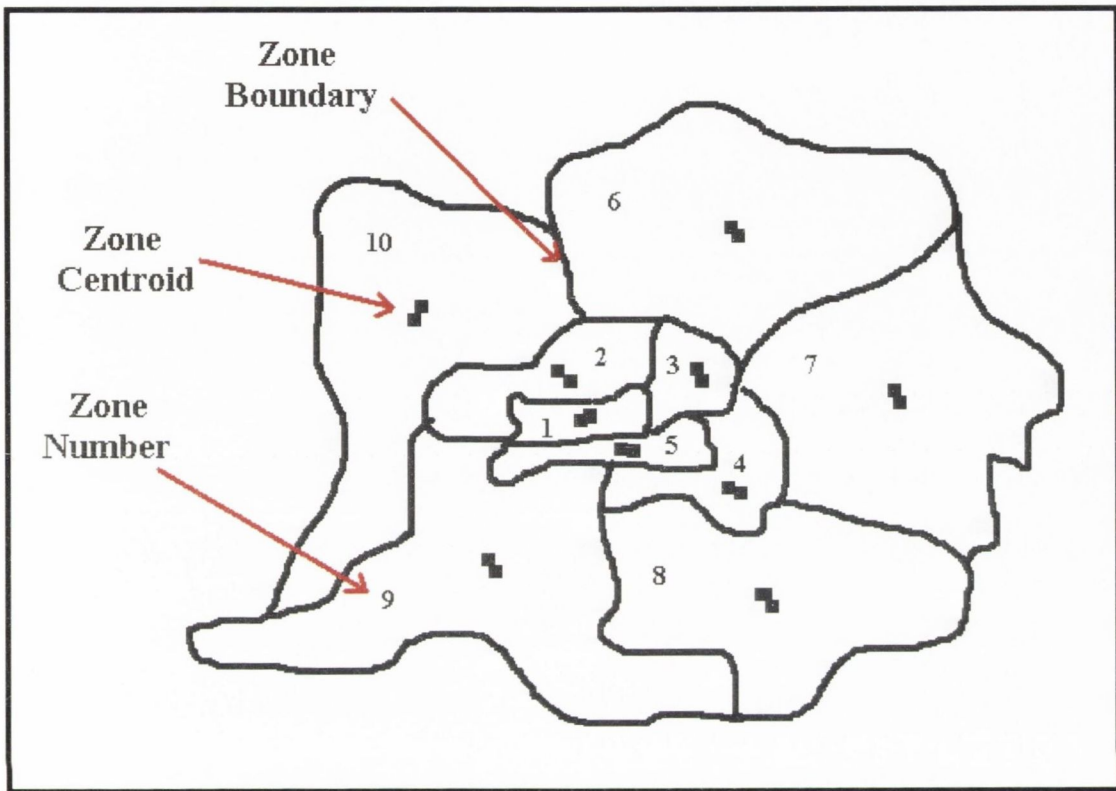
The area of interest is defined as:

*“The area within which trip patterns will be significantly affected by the implementation of transport proposals.”* (Byrne, 2002)

Once this area has been identified then an analysis of the existing trip patterns is necessary plus a prediction of future year trip patterns, as was the case in 1997 when the DTO updated trip patterns on the existing 1996 DTO Traffic Model. The DTO Traffic Model is broken up into a hierarchical zoning system with zone boundaries, centroids and numbers; see Figure 4.2 for a typical zoning system with:

- 432 fine zones, with finer zones in the City Centre and coarser zones further out, allowing for a combination of robust forecasting of socio-economic and travel activity variables with the detail required for accurate mode choice and assignment;
- 367 zones based on District Electoral Divisions (DEDs), which are used in association with planning data;
- 58 strategic zones like that of the original DTNM; and,
- 20 coarse zones.

Origin and destination zones may be considered as a subset of the nodes known as centroids, see Section 4.2.3.



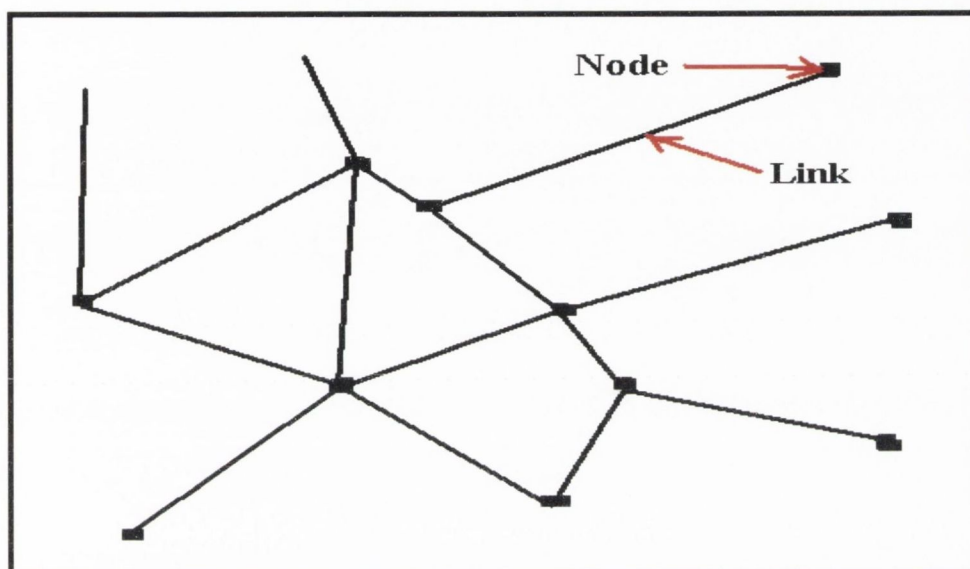
**Figure 4.2:** A typical zoning system with zone boundaries, centroids and numbers.

#### 4.2.3 Transportation networks

The supply of transport infrastructure in the GDA is represented by a series of networks. Each mechanised mode has its own network; one representing the road system for use by private cars and commercial vehicles known as the highway network, and those representing the bus system and rail services known as public transportation networks.

Transportation networks are defined in terms of links and nodes. In the highway network links are defined as the road sections between junctions and nodes are defined as the junctions. In the rail network nodes are train stations and links are the track sections between them. Public transport is defined in terms of speed, efficiency, frequency and routing through nodes. Figure 4.3 shows a basic sample of a transportation network.

The bus network used to be modelled as part of the highway network with full Dublin Bus routings, time and operating schedules up until the 2002 DTO model update.

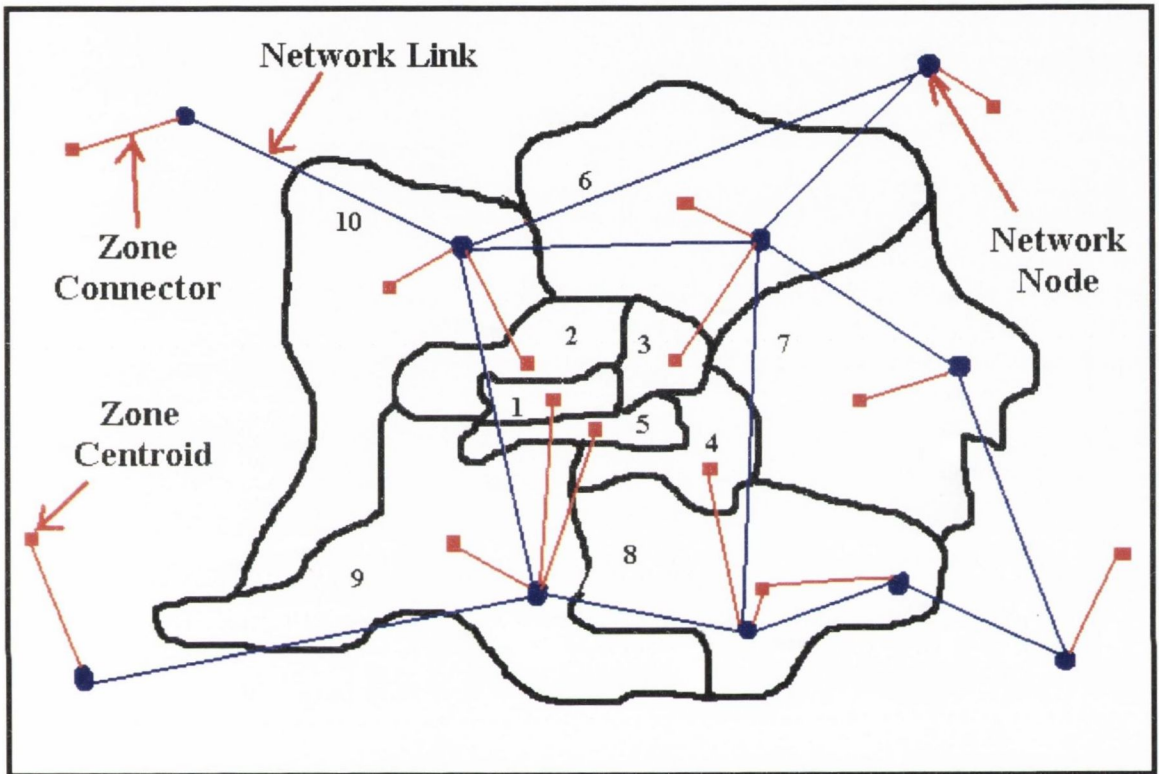


**Figure 4.3: A transport network with nodes and links.**

However, since 2002 a new model developed by MVA Consultants known as TRIPS, or Transport Improvement Planning System (see <http://www.citilabs.com/trips/index.html> for more details), has been employed to model the public transportation networks, taking over from the less proficient SATCHMO (Saturn Travel Choice Model) extension to the SATURN Suite, which is also designed to assess public transportation networks. All suburban rail is modelled as part of the rail network and incorporates the full Irish Rail time schedules and operating frequencies for services. A diagram of how the zones are connected with the transport network using centroid connectors can be seen in Figure 4.4. Centroid connectors may be considered a subset of the links.

The highway network is made up of a simulation network and a buffer network, combining to give 1,220 simulation junctions and 3,900 links throughout the GDA. The simulation and buffer networks represent two different levels of detail to reflect the complexity of congestion phenomena in the GDA. Links in the buffer network use a simplified representation of delay as a function of the flow on the link itself. Links in the simulation area, which represents the inner city and urban parts of the GDA, i.e., the Metropolitan Area, consider delays at junctions as a function of traffic control characteristics and of the flows on all approaches.

Simulation links are especially useful where congestion is likely to be significant because of the detail in which they are modelled. Separate transportation networks have



**Figure 4.4:** A transport network with zones, centroids, nodes, links and centroid connectors.

been produced at peak and off-peak periods due to the different traffic signal timings in the two periods. The rail and bus networks for the peak and off-peak models represent the different service patterns for those times of the day.

#### 4.2.4 The transportation model structure: classic 4 - stage model

The classic 4-stage model used by the DTO, see Figure 4.5, is the same as that used by the DTI, as described in Chapter 2 Section 2.4.4. However, the description in Section 2.4.4 is somewhat theoretical and what follows is a more practical description of how the theory applies to the DTO Traffic Model.

#### 4.2.4.1 Trip generation – attraction stage

There are two aspects to trip generation; generation and attraction. In the DTO Traffic Model the number of trips generated in a particular zone is a function of the socio-economic properties of the zone while the trips attracted to a zone are a function of the land use properties of the zone. Trip generation and attraction can be stated as follows:

$$O_i = f(x_{i1}, x_{i2}, x_{i3}, \dots) \quad \text{Equation 4.1}$$

and,

$$D_j = f(y_{j1}, y_{j2}, y_{j3}, \dots) \quad \text{Equation 4.2}$$

where:

- $O_i$  are the trips generated by zone 'i' and  $D_j$  are the trips attracted to zone 'j'; and,
- $X_i$  are the socio-economic properties of zone 'i' while  $Y_j$  are the land use properties of zone 'j'.

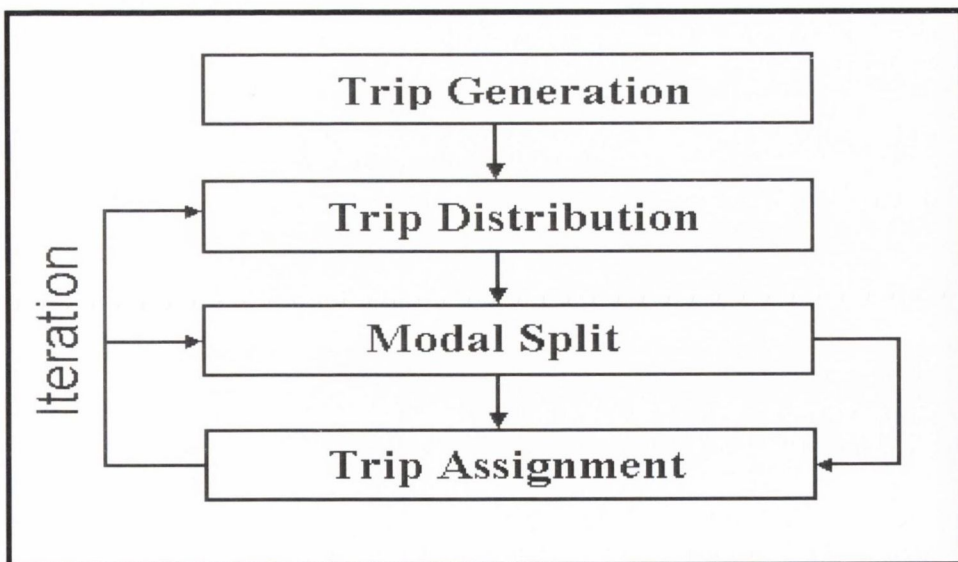


Figure 4.5: Four-stage model used in the DTI's DTNM and in the DTO Traffic Model.

#### 4.2.4.2 Trip distribution stage

Trip distribution models the distribution of trips generated by each origin zone that are attracted to each destination zone. The output of the process is a matrix of trip distributions for origin zone 'i' to destination zone 'j', i.e.,  $T_{ij}$ , such that:

$$\sum_i T_{ij} = D_j \quad \text{Equation 4.3}$$

and,

$$\sum_j T_{ij} = O_i \quad \text{Equation 4.4}$$

The base year trip generation and distribution were derived from data obtained in the same multi-modal origin / destination surveys described in Chapter 2 Section 2.4.4 for the DTI's DTNM and updated by the DTO's disaggregate stated preference surveys for the DTO Traffic Model. Trip end totals were validated against demographic and land use data while the trip demand forecasts were based on demographic, land use and macro-economic forecasts, and, on the expansion of base year trip rates.

#### 4.2.4.3 Modal split stage

Determination of the choice of mode for a trip is known as modal split. The most important modal split in the DTO Traffic Model is the split between public and private transportation. After this, a further split occurs between the different modes of public transportation. This split is determined by assigning a generalised cost to each mode of travel, derived from disaggregate stated preference surveys, and then, in the case of the DTO Traffic Model, using a hierarchical logit model to determine the split between public and private transportation and then between rail and bus transportation. This logit model is described in Chapter 2, Section 2.4.4. However, while the equation for the generalised cost of travel by public transport has remained the same, see Section 2.4.4, for the updated DTO Traffic Model, the generalised cost of travel by private car has been abbreviated to the following:

$$C_{ij} = a_1V_t + a_2V_d + a_3P + a_4 \quad \text{Equation 4.5}$$

where;

- $C_{ij}$  is the generalised cost of travel from zone 'i' to zone 'j';
- $V_t$  is the in-vehicle time, including time spent looking for parking;
- $V_d$  is the in-vehicle distance, including distance spent looking for parking;
- $P$  is the parking charge; and,
- $a_4$  is a constant representing other fixed costs, e.g., tolls, congestion charging, etc.

Using model calibration suitable values for  $a_1$ ,  $a_2$ ,  $a_3$ , etc. can be found for the expressions of generalised cost for both public and private transportation, i.e., rail, bus and private car. This allows an expression to be found for the generalised cost of travel between any two zones, 'i' and 'j', for any given mode of travel.

#### 4.2.4.4 Assignment stage

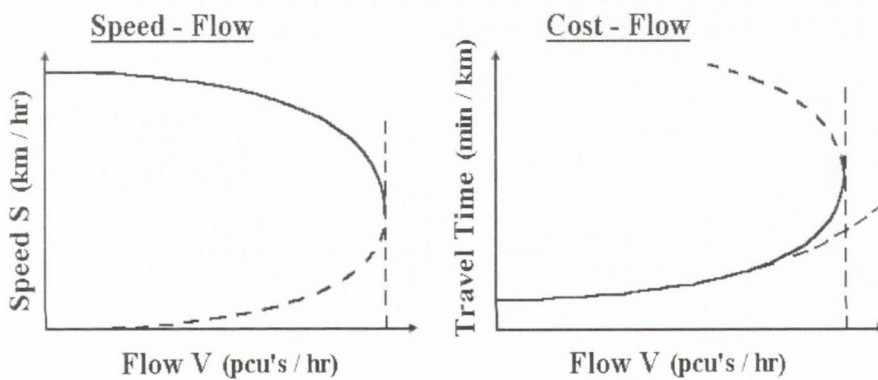
Trip assignment is the last stage of the classic 4-stage model. Mode choice and assignment in congested urban networks are not independent of each other. This is explicitly recognised in the DTO Traffic Model, e.g., trip makers who have a private car at their disposal but may or may not still use public transportation are subject to 'equilibration' between mode choice and assignment models. This means that there is consistency in the treatment of generalised cost, the key determinant of travel choice, between assignment and mode choice.

After the modal split each trip distribution matrix is assigned to the appropriate transportation network. This assignment is calculated based on the minimum generalised cost of travel between each 'ij' pair of zones. In the case of highway trip assignment, i.e., the private car distribution matrix, the time taken to complete a journey



on a given link is a function of the characteristics of the link and the volume of traffic on it. These characteristics include the link distance, link free flow speed, link capacity speed, link capacity and the link speed-flow relationship.

For highway links that do not have junctions the speed on the link can be related to the traffic flow on the link by a speed flow curve as seen in Figure 4.6. If the speed flow curve is inverted then a cost-flow curve is produced, also Figure 4.6, which allows the link journey times to be calculated for links on a highway network.



**Figure 4.6: An example of Speed versus Flow and Cost, measured in travel time, versus Flow for a sample link representing a section of highway.**

The trips are assigned to the highway network based on the Wardrop Equilibrium Principle (Wardrop, 1952). The Wardrop Equilibrium Principle states that:

*“At equilibrium all routes used by a particular user class are minimum cost routes as defined by that user class while all other (permitted) routes have equal or greater cost.”*

However, if all trip makers perceive their costs in the same way then the Wardrop Principle can be re-written as follows:

*“Under equilibrium conditions traffic arranges itself in congested networks such that all routes between any Origin / Destination pair have equal and minimum costs, while all unused routes have greater or equal costs.”*

To achieve Wardrop's Principle over a network, in practical terms, involves an iterative process, which gradually converges until equilibrium conditions, or conditions acceptably close to equilibrium, have been obtained. The SATURN traffic simulation and assignment model is used to carry out this procedure for highway networks in the DTO Traffic Model. Each iteration of the assignment process has four basic stages:

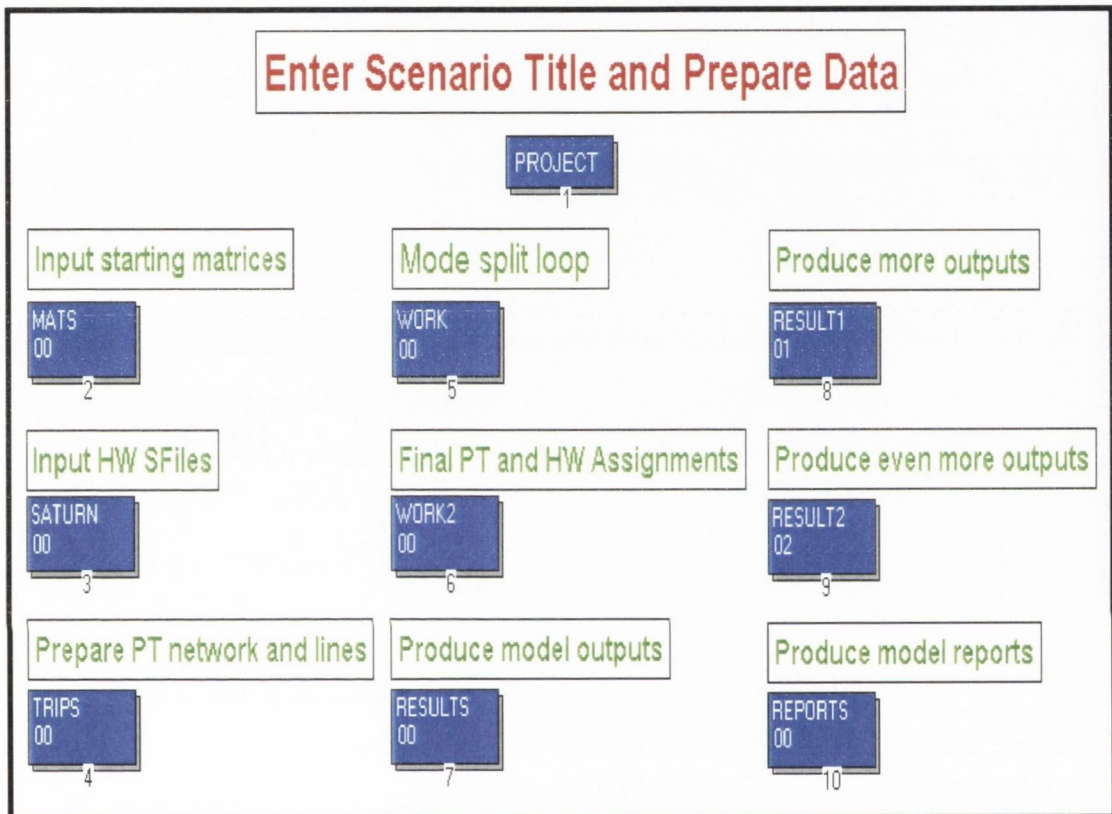
- Completing an all-or-nothing assignment on the cheapest route;
- Calculating the flows as a linear function of the previous flow plus the flow from the all-or-nothing assignment;
- Re-calculating the travel costs using these new flows for each link; and,
- Repeating the three steps above until the travel costs on each link are approximately equal, i.e., within a set convergence parameter, and the Wardrop equilibrium conditions have been reached.

Since the 2002 update public transport assignment has been carried out by TRIPS instead of SATCHMO. TRIPS is more flexible than SATCHMO allowing trip transfer between bus and rail services at intersection points on the public transportation network. TRIPS has a number of other advantages. For instance, SATCHMO calculates the costs associated with public transportation from the first assignment iteration and uses these costs to ascertain the delays for buses as they travel through a congested network. Whereas TRIPS takes a set of costs at the end of every new assignment iteration giving a more accurate estimation of the delay costs for the movement of buses at the end of the model run. TRIPS also allows more freedom in that one can set different fare tables for different public transport modes, include other penalties such as crowding penalties, and, assess 'Park and Ride' more accurately by calculating the combined cost of driving to the site of a public transport service and then travelling on it.

#### *4.2.5 Overview of the DTO Traffic Model layout*

The software platform adopted in the DTNM was SATURN for highway modelling and SATCHMO for public transportation modelling. Since the 2002 update of the DTO

Traffic Model the SATURN Suite continues to be used for modelling highway transportation while public transportation is, for the most part, assessed using TRIPS.



**Figure 4.7: Overview of the DTO Traffic Model from Data Preparation (1) to Results (10) including the use of SATURN and TRIPS.**

While TRIPS has a number of major advantages over the SATCHMO package there is one major drawback. Within the DTO Traffic Model, if SATURN is used in conjunction with TRIPS, see Figure 4.7, then the computing time required is four or five times greater, depending on the values or involvement of certain parameters, than that of the SATURN – SATCHMO combination, which takes between six and eight hours to run (Byrne, 2002).

This research is looking at means of improving the congestion difficulties on the highway network of the GDA using demand management measures. One of the main problems with the use of the DTO Traffic Model when applied to the highway network, which requires several model runs, is that each model run takes between six and eight

hours to complete. It is proposed that high performance computing will reduce this run time, significantly increasing the scope for research in the area of demand management and allowing an assessment of transportation demand management proposals, and in particular congestion charging options, in the GDA.

### **4.3 *The SATURN traffic simulation and assignment model***

#### **4.3.1 *Introduction***

SATURN is a suite of transport network analysis programs developed at the Institute for Transport Studies (ITS), University of Leeds by Dr. Dirck van Vliet and distributed by WS Atkins of Epsom since 1981. The model, version 4.0, was first described by Hall et al. (April 1980). It was initially designed for “self contained areas in large cities or for free standing towns of less than 100,000 population” to achieve reasonable computer times (Hall et al., 1980). The article details the first practical application on the town of Harrogate in North Yorkshire.

In 1982 an updated article was published describing a “considerably improved and extended” version having been applied in a number of studies including a traffic management study in Liverpool (Choraffa et al., 1982 and 1983) and studies in Sydney, Australia, and Auckland, New Zealand. Testing was done in central Manchester involving SATURN, version 8.0, in 1987 to validate the SATURN and ME2 (Matrix Estimation from Maximum Entropy) models, for a network before and after the introduction of a pedestrianisation scheme (Matzoros et al., 1987). It was found that both models were accurate to within approximately 12%. SATURN, version 9.4, now has over 250 users in approximately 30 countries world-wide and has been tested and applied by consultants, universities and planning authorities for the evaluation of traffic management and demand management schemes (ITS web site, 2003).

A number of road pricing studies involving SATURN in more recent times include ROCOL, The Review of Charging Options for London (GOL, 2000), four studies conducted in different Scottish towns by Oscar Faber in 1995 ([www.scotland.gov.uk](http://www.scotland.gov.uk))

and a number of other studies across the UK including those of Santos (2000) and May et al. (2000).

The model has five main functions (ITS and WS Atkins, 1998), which are:

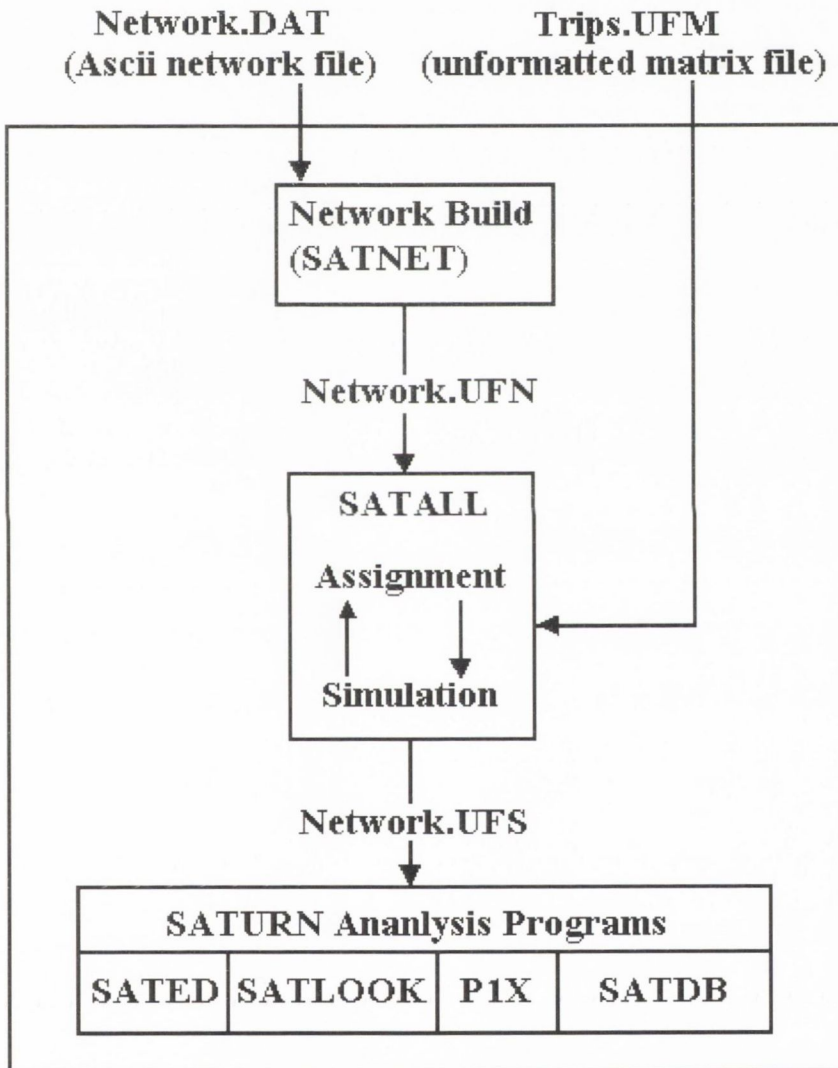
- (1) The analysis of traffic management schemes over localised networks, of the order of between 100 and 1,000 junctions on average, as a combined traffic simulation and assignment model;
- (2) The analysis of much larger networks of between 6,000, for the standard PC version, and 30,000 links as a conventional traffic assignment model;
- (3) The analysis and modelling of individual junctions using the simulation model;
- (4) The analysis of transport networks and the production and manipulation of transportation network databases; and,
- (5) The production and manipulation of trip matrices with the use of a matrix package.

There are two types of input to the SATURN model including a specification of the demand for travel defined by an origin/destination trip matrix, time profile, etc. and a specification of the supply provided by the transportation network, which is defined using links, nodes, link lengths, capacities, banned turns etc. The output consists of link flows, route choices, turning movements, etc., given by the traffic pattern produced by the set of rules which define how trip makers select their mode and route of choice (van Vliet, 1975).

#### 4.3.2 *The SATURN model framework*

The basic SATURN model has seven main programs: the first three are concerned with assigning traffic flows to the transportation network; and, the last four are concerned

with the analysis of loaded transportation networks, where the trip matrix has been assigned to the transportation network, i.e., the demand has been assigned to the supply (van Vliet and Hall, 1998). Figure 4.8 shows the SATURN model framework with SATALL being used in place of SATEASY and SATSIM, which are described below.



**Figure 4.8: The SATURN simulation and assignment model framework showing SATALL in place of SATEASY and SATSIM.**

A further program for the production and manipulation of trip matrices, known as MX, is also used quite frequently. MX is used to input all the necessary information about a trip matrix from a number of sources and to create the input file, 'Trips.UFM'. This file is an *UnFormatted Matrix* (.UFM) file containing all the necessary data that the

SATURN Suite needs to use in order to manipulate and edit the newly created trip matrix in further programs, particularly in the simulation and assignment processes associated with SATALL.

The programs that deal with loading traffic, in other words, the simulation and assignment of traffic onto the transportation network, are described as follows:

- The network build program, SATNET, which takes the input ASCII data files and prepares, or reformats, the information as an '*UnFormatted Network*' file (Network.UFN) so that it can be used as input into the next three programs;
- The assignment program, SATEASY, which assigns traffic to the transportation network on the basis of delays, either measured through chronometric or monetary means given as output by the simulation, and produces an '*UnFormatted Assignment*' file (Assignment.UFA);
- The simulation program, SATSIM, which takes the transportation network with its associated assigned trips from the assignment stage and models the passage of traffic through the transportation network outputting the resulting delays from congestion and other related information in an '*UnFormatted Simulation*' file (Simulation.UFS); and,
- The combined simulation and assignment program SATALL, which combines SATEASY and SATSIM into a single program switching automatically between the two processes until a convergence parameter has been reached. Once this is the case then an '*UnFormatted SATALL*' file (SATALL.UFS) is output for use with the analysis programs described below.

The SATALL model either iterates between the processes of assignment (SATEASY) and simulation (SATSIM) manually or iterates through them automatically (SATALL) until Wardrop's Equilibrium has been satisfied.

The analysis and display programs for assessing loaded transportation networks are described as follows:

- The network plot program, PIX, displays link-based data either to a terminal or as line printer output for the production of hard copy. It also contains virtually all the functions contained in the three programs that follow;
- The analysis program, SATLOOK, allows the user to assess specific model results within the context of a large output file and print out a description of traffic conditions from model results;
- The network editing program, SATED, allows nodes, usually junctions, to be edited and simulated on an individual basis using interactive commands; and,
- The data base analysis program, SATDB, allows a free-format manipulation and display of data.

The analysis programs PIX, SATLOOK, SATED and SATDB, enable users to examine output from the assignment, simulation or combined assignment-simulation modelling processes. The analysis programs themselves are of little interest in the re-programming of the sequential SATURN model and they will not be discussed further because they are predominantly for analysis purposes only and do not affect the computationally intensive modelling procedure.

#### **4.4 Initial analysis of SATURN**

##### *4.4.1 Introduction*

Despite the extent to which the Dublin Transportation Initiative (DTI, 1995) strategy has been implemented and the substantial investment still planned for Dublin City and its hinterland under the Initiative, it is apparent that traffic congestion will remain a serious problem in the Greater Dublin Area (GDA) for some time to come. As a result, there remains the need to assess means of analysing the problem of traffic congestion in



order to find solutions. The DTO Traffic Model, which uses the SATURN traffic assignment and simulation model, is used to analyse potential transport planning solutions over the GDA (DTO, 1998).

One of the objectives of this research has been to assess the SATURN model (van Vliet, 1982), as applied to the DTO Traffic Model, with a view to finding the areas within the model that take up the most Central Processing Unit (CPU) time, and to reprogram these areas to run in parallel. Reprogramming the most time consuming, or computationally intensive, areas of the model in parallel will allow for a more efficient use of computer time over a variety of network configurations and sizes, and in doing so, will result in an overall improvement in efficiency for the DTO Traffic Model. The success of this overall objective will be measured through the achievement of three lesser objectives summarised as the following:

- To produce output from the reprogrammed parallel model that will demonstrate its ability to converge as quickly as the sequential model;
- To produce output that will demonstrate the accuracy of the output data from the new parallel model as compared with output from the sequential (standard format) version of the SATURN model using the same input data; and,
- To produce output that will compare the performance of the new model against that of the sequential SATURN model and to show the improvement of the performance over that of the sequential model for the same test scenarios.

This Section analyses the model to assess if there are areas that take up large quantities of CPU time and what the potential is for reprogramming any such areas using parallel programming techniques and hardware.

#### *4.4.2 Initial analysis*

The assignment and simulation section, as opposed to the analysis and evaluation section, of the SATURN model has three different parts: the input, the model and the output. The input and output are once-only transfers of data and results to and from an

internal format, known as 'unformatted' in SATURN, and will not be considered further. However, the processes involved represent a not inconsiderable overhead from a parallelisation standpoint and this will be considered at a later stage in the thesis.

It was found by Greenwood and Taylor (1993) that for the CONTRAM (CONTINUOUS TRAFFIC ASSIGNMENT MODE) model (Leonard and Gower, 1982; Leonard et al., 1989; Taylor, 1990) most of the computationally intensive areas, the areas where the Central Processing Unit (CPU) spends most of its time, are within the assignment area. CONTRAM is a dynamic traffic assignment model, which models time-variable demands, routes and network variables (Taylor, 1992), that was developed by the UK Transport Research Laboratory. As a starting point for the assessment of SATURN it was initially assumed for this research that because of the similarities between the CONTRAM and the SATURN models, similar results would be obtained for the CPU time spent within the assignment area.

The computationally intensive part of the SATURN model is the SATALL program. SATALL deals with loading the traffic onto the transportation network. SATALL is made up of an assignment and a simulation. The assignment section assigns traffic on the basis of the delays given by the simulation. The simulation section models in detail the passage of traffic through the network and the resulting delays. The program SATALL loops between the assignment and simulation until equilibrium is reached. The SATALL program was profiled using a piece of software called '*Gprof*', with the optimiser switched on, run on the UNIX operating system to show on what and for how long the CPU spends its time. These results can be seen in Table 4.1.

Table 4.1 gives a detailed profile of the top 38 most time consuming routines relative to the overall running time, categorised by the percentage of CPU time used while running the computationally intensive SATALL program, when applied to a sample problem supplied by the DTO. The details of these routines are unimportant, as the computational time distribution between them is what is significant (Greenwood and Taylor, 1993). The second column, '% (Time)', gives the percentage of CPU time that each routine uses while SATALL runs through a given problem. The next column,

'Cumulative Time (Secs)', gives the cumulative time used up by successively less time consuming routines in the SATALL program as the assignment/simulation iterates to convergence. Column four, 'Self (Secs)', gives the time that each routine uses per iteration of itself, as opposed to the additional time spent in subroutines that happen to be called from that routine. The column, 'Self (No. of Calls)', gives the number of times that each routine is called from another routine while the 'Time (ms/call)' and 'Total Time (ms/call)' columns give the milliseconds used within each routine per iteration of itself, i.e., per call, and the total milliseconds used per call including the time spent in any subroutines that happen to be called from that routine, respectively.

Table 4.1 shows that the nested subroutine used to solve the shortest path problem, '*desopo\_cb*', demands over 80% of the overall CPU running time for the SATALL program. As well as this, the subroutine, '*loadit\_plus*', (used to load, or assign, the trips onto these shortest paths) that calls the nested shortest path subroutine, '*desopo\_cb*', uses an additional 5% of the total CPU running time. This only leaves approximately 12% of the overall running time for all of the other routines that make up the program SATALL. The details of many of these subroutines are not shown in Table 4.1 as there are too many of them to list but 99.1% of the total time for the program is accounted for by the 38 routines shown.

#### 4.4.3 Conclusion

- The processes involved in the assignment area of the SATURN model were found to have the most potential for performance improvement using parallel processing techniques.
- Since both the '*desopo\_cb*' (shortest path) and the '*loadit\_plus*' (trip assignment) routines are both important sections of the assignment, and because these routines give the assignment over 85% of the overall CPU running time for the execution of SATALL including input and output, it was decided that the assignment will be targeted for further research into potential time saving development using high performance computers.

**Table 4.1: Profile of CPU related figures and number counts for a number of the routines used in the execution of the SATALL program.**

Subroutines used in SATALL program	% (Time)	Cumulative Time (Secs)	Self (Secs)	Self (No. of Calls)	Time (ms/call)	Total Time (ms/call)
.desopo_cb [7]	80.8	886.82	886.82	83711	10.59	10.69
.loadit_plus [6]	4.9	940.57	53.75	194	277.06	4897.06
._mcount [16]	1.9	961.52	20.95	NG	NG	NG
.desopo_cb_2[17]	1.9	981.85	20.33	1726	11.78	11.88
.tommix[15]	1.4	997.01	15.16	200889	0.08	0.11
.init[20]	1.2	1009.9	12.89	337577	0.04	0.04
.initr[23]	1	1020.93	11.03	490193	0.02	0.02
.opatt[21]	0.6	1027.24	6.31	242663	0.03	0.05
._pow[18]	0.5	1032.52	5.28	10094610	0	0
.saxvt[22]	0.4	1037.45	4.93	6807060	0	0
.loginner2[27]	0.4	1042.06	4.61	8479574	0	0
.expinner2[28]	0.4	1046.47	4.41	8479574	0	0
.whasig[29]	0.3	1050.04	3.57	4113402	0	0
.movset[26]	0.3	1053.25	3.21	4603361	0	0
.itdset[30]	0.3	1056.21	2.96	7493280	0	0
.gapx[19]	0.2	1058.67	2.46	425885	0.01	0.03
.memcpy[33]	0.2	1060.98	2.31	NG	NG	NG
.totflo[35]	0.2	1063.22	2.24	7652970	0	0
.setacc[11]	0.2	1065.35	2.13	242663	0.01	0.22
._stack_pointer[36]	0.2	1067.44	2.09	NG	NG	NG
.mixlin[12]	0.2	1069.22	1.78	154805	0.01	0.22
.cijadd[39]	0.1	1070.59	1.37	7135337	0	0
.lovset[41]	0.1	1071.94	1.35	5905959	0	0
.loadit_plus_2[14]	0.1	1073.21	1.27	4	317.5	5448.76
.pldis[44]	0.1	1074.44	1.23	185793	0.01	0.01
.qincrement1[46]	0.1	1075.51	1.07	NG	NG	NG
.prisig[10]	0.1	1076.56	1.05	60568	0.02	1.14
.qincrement[47]	0.1	1077.55	0.99	NG	NG	NG
.llset[49]	0.1	1078.52	0.97	1950350	0	0
.mixriv[50]	0.1	1079.49	0.97	154805	0.01	0.01
.mucfw[5]	0.1	1080.43	0.94	2	470	482074.8
.set_flow_by_lane[34]	0.1	1081.35	0.92	412572	0	0.01
._log[53]	0.1	1082.17	0.82	1078977	0	0
.delay_cfp[51]	0.1	1082.99	0.82	55214	0.01	0.02
.saobjt[38]	0.1	1083.8	0.81	98	8.27	14.67
._xlfBeginIO[56]	0.1	1084.56	0.76	NG	NG	NG
.rbout[37]	0.1	1085.31	0.75	2911	0.26	0.52
.outin[31]	0.1	1086.05	0.74	63551	0.01	0.05

- Hislop et al. (1991) also identified that, amongst others including CONTRAM, the equilibrium assignment model SATURN is suitable for implementing on parallel computers.

## **4.5 Trip Assignment**

### *4.5.1 Introduction*

This research concentrates on the standard default assignment procedures used by SATURN. These can be found in the SATURN User Manual, see van Vliet and Hall (1998). These procedures are also the standard network requirements and assignment procedures used by the DTO for most of the work that it carries out using the SATURN model.

The principle of road assignment was originally formulated by Wardrop in 1952, see Section 4.4.3 below. The assignment procedure accepts as input a trip matrix and assigns those elements to minimum cost routes through the network based on the current cost - flow relationships, described in Section 4.2.4, either user-input or derived from the simulation. There are various methods of assignment available. In the case of SATURN the user has a choice of two standard methods: Wardrop Equilibrium Assignment and Stochastic User Equilibrium Assignment. SATURN also allows for Multiple User Class Assignment and Elastic Equilibrium Assignment, see Ortuzar and Willumsen (1994). The relevant areas in relation to the DTO Traffic Model have been summarised briefly in the sections that follow.

The DTO uses Wardrop or User Equilibrium Assignment over multiple user classes for most of its work. It is proposed that the new high performance computing version of the SATURN model will be developed to analyse these procedures more efficiently, and in doing so, allow for more efficient analysis of potential congestion charging schemes in the GDA using the DTO Traffic Model.

#### 4.5.2 Multiple User Classes

Multiple user classes refers to the consideration of more than one means of transportation over a given transportation network at the same time, e.g. considering the effect of public transportation and of private cars over a highway network at the same time. The interaction between different user classes can be based either on a Wardrop Equilibrium Principle (Wardrop, 1952) or on a Stochastic Equilibrium Principle (Ortuzar and Willumsen, 1994). The Stochastic principle states that:

*“At equilibrium individuals within each user class perceive all routes used as being minimum cost routes (as well as being permitted routes).”*

While the Wardrop principle, described in Section 4.2.4 above, states that:

*“At equilibrium all routes used by a particular user class are minimum cost routes as defined by that user class while all other (permitted) routes have equal or greater cost.”*

#### 4.5.3 User Equilibrium Assignment

The default assignment procedure within SATURN is based on Wardrop’s Principle of Traffic Equilibrium (Wardrop, 1952), which states that:

*“Traffic arranges itself on congested networks such that all routes used between any Origin and any Destination have equal and minimum costs while all unused routes have equal or greater cost.”*

This principle is based on two assumptions:

- (1) That each user chooses their route based on the minimum perceived cost; and,
- (2) That all users perceive their costs in the same fashion.

While the first assumption seems like a representation of realistic driver behaviour, the second assumption is weaker as it is well known that not all drivers perceive their costs in exactly the same way and that they choose different routes for a variety of different

reasons. Unlike certain stochastic methods of assignment that may not converge if the network being modelled is congested (van Vliet, 1979) equilibrium methods based on Wardrop's Principle have led to models which are stable on heavily congested transportation networks and, to algorithms that converge to an equilibrium solution (Hislop et al., 1991). An example of which is the Frank-Wolfe algorithm (Frank et al., 1956) described in Section 4.5.4 below.

Ortuzar and Willumsen (1994) make reference to work carried out by Beckman et al. (1956) where the set of flows,  $V_a$ , satisfying Wardrop's Principle can also be obtained by finding the set of flows that minimise a certain "*objective function*" subject to all the path flows that make up the loaded trip matrix being positive, e.g., for a single user class:

$$Z = \sum_a \int_0^{V_a} C_a(V) dV \quad \text{Equation 4.6}$$

where;

- $C_a$  is equal to the cost of travelling from link '0' to link 'a', and,
- $V_a$  is equal to the flow over those links.

This is extremely useful in that it enables one to establish algorithms that, by minimising the objective function 'Z', guarantee finding an equilibrium solution. The objective function for a multiple user class assignment under Wardrop Equilibrium conditions may look like the following equation given here (Ortuzar and Willumsen, 1994):

$$Z = \sum_a \int_0^{V_a} t_a(V) dV + \sum_0^a \sum_0^k \sum_0^m b'k.m \ d_{a.k} \ V_{a.m} \quad \text{Equation 4.7}$$

where;

- $t_a$  is equal to the travel time from link '0' to link 'a' and ' $V_a$ ' is equal to the flow over those links;
- $b'k.m$  is equal to the value of one unit of fixed cost component 'k', which can be measured in time or distance or a combination of a number of factors, to user class 'm' divided by the value of time as perceived by user class 'm';
- $d_{a.k}$  is equal to the k'th fixed cost component on link 'a'; and,
- $V_{a.m}$  is equal to the flow over link 'a' by user class 'm'.

There are a number of algorithms that can be used to achieve Wardrop Equilibrium within the SATURN Suite. These include the Frank-Wolfe algorithm (Frank and Wolfe, 1956), the PARTAN, or **PAR**allel **TAN**gent algorithm (Florian et al., 1983), and, the Method of Successive Averages as described by Ortuzar and Willumsen (1994); all of which are briefly described in the sections that follow.

#### 4.5.4 The Frank-Wolfe Algorithm

The adaptation of the Frank-Wolfe algorithm, as used within the SATURN model, is described as follows:

- (1) Assign all trips to Origin-Destination (O-D) distribution paths to produce an initial set of link flows,  $V_a^{(n)}$ , where  $n = 1$ . Conventionally, the first assignment is an all-or-nothing assignment with the initial set of link costs measured in time units and set to their "free-flow" values.
- (2) Create new link times by altering the old link times in accord with the current flows,  $V_a^{(n)}$ ; i.e., set:  $C_a^{(n)} = C_a(V_a^{(n)})$ .
- (3) Build a new set of shortest paths based on the new link costs,  $C_a^{(n)}$ , and assign all trips,  $T_{ij}$ , to them to produce a set of "auxiliary" all-or-nothing flows,  $F_a^{(n)}$ .
- (4) Generate an "improved" set of link flows,  $V_a^{(n+1)}$ , as a linear combination of the old and the auxiliary flows, i.e:



$$V_a^{(n+1)} = (1 - \lambda) * V_a^{(n)} + \lambda * F_a^{(n)} \quad \text{Equation 4.8}$$

where;

- ‘ $\lambda$ ’ is chosen so that the “new” flows,  $V_a^{(n+1)}$ , minimise the objective function, ‘Z’, to satisfy Wardrop’s Equilibrium as described in the sections above.
- (5) Return to step (2) unless some convergence criterion has been satisfied; for example, the maximum number of loops as specified by the user defined parameter, NITA, has been exceeded. NITA is a parameter within SATURN defining the maximum number of assignment iterations to be completed.

Each trip from an origin to a destination is represented by an O-D pair, ‘ij’. At each stage of the algorithm a new set of routes is calculated for all origin-destination pairs, ‘ij’, and then a certain proportion of all previously assigned trips are shifted onto these routes. Hence, after say five iterations, there will be up to five different routes for each ‘ij’ pair (allowing for the same route to be chosen more than once) with certain fixed proportions of trips assigned to each creating a trip distribution over the network for each ‘ij’ pair.

Starting from an initial feasible solution, i.e., trips assigned to links based on their free-flow values, the method generates a feasible direction by linearising the objective function. This means solving a linear (Taylor series expansion) programming sub-problem and then finding an improved solution on the line segment between the current solution and the solution of the sub-problem.

At each step in an iteration there is a feasible solution and the algorithm takes the latest all-or-nothing assignment to provide a direction in which to take in order to achieve this solution. This direction can be seen as an approximation to minimising the objective function, ‘Z’. The Frank-Wolfe algorithm then seeks to find a second feasible direction using a linear approximation to the objective function. To minimise the approximation to the solution of the function, ‘Z’, the algorithm chooses routes such that their

corresponding costs are minimised. In this way the algorithm fulfils the requirements of Wardrop's Equilibrium principle by only choosing a new route such that it has either less costs or equal costs than the previous route. In other words, all routes with assigned trips are minimum cost routes with all other routes having equal or greater costs. There are numerous descriptions of the Frank-Wolfe algorithm including those in Daganzo (1979) and LeBlanc (1975) amongst others.

#### 4.5.5 PARTAN assignment

The PARTAN, or **PAR**allel **TAN**gent, algorithm is a variant of the Frank-Wolfe algorithm for solving Wardrop equilibrium with a single user class. It introduces a new step into the Frank-Wolfe algorithm, described in Section 4.5.4, between steps 4 and 5 of the algorithm. This new step combines the flows,  $V_a^{n+1}$ , at the end of step 4 with the second previous solution for iterations greater than 2, i.e:

$$V_a^{(n+1)} = \left(1 - \delta^n\right) * V_a^{(n+1)} + \delta^n * V_a^{(n-1)} \quad \text{Equation 4.9}$$

where;

- $\delta^n$  is chosen to minimise the objective function,  $Z(V_a^{(n+1)})$ .

LeBlanc et al. (1985) and Florian et al. (1987) have described the PARTAN method as applied to equilibrium assignment and they have shown that it has improved convergence over the Frank-Wolfe algorithm on its own. Arezki et al. (1990) have also provided an analytical implementation of the PARTAN variant to the Frank-Wolfe algorithm for user equilibrium assignment.

#### 4.5.6 Method of Successive Averages (MSA)

MSA is an iterative assignment algorithm. It is used with stochastic user equilibrium (SUE) assignment. The difference between stochastic and Wardrop's equilibrium is that in SUE models each driver is meant to define travel costs individually instead of using a single definition of costs applicable to all drivers (Ortuzar and Willumsen,

1994). In other words “*no-one can improve his/her perceived travel cost by unilaterally changing route*” (Hislop et al., 1991). This is a more realistic approach in terms of modelling behaviour than the assumptions associated with Wardrop’s Principle of Traffic Equilibrium.

The flow on a link is calculated as a combination of an auxiliary all-or-nothing assignment in the present iteration and the current flow on the previous iteration. The algorithm is described as follows:

- (1) Initialise all flows  $V_a = 0$  and make  $n = 0$ . Conventionally, link costs are set to their “*free-flow*” values.
- (2) Build a set of minimum cost paths with the current costs and make  $n = n + 1$ .
- (3) Load the matrix onto these paths obtaining a set of auxiliary flows,  $F_a$ .
- (4) Calculate the current flows as:

$$V_a^{(n)} = (1 - \lambda) * V_a^{(n-1)} + \lambda * F_a \quad \text{Equation 4.10}$$

with,  $0 \leq \lambda \leq 1$

Then calculate a new set of current link costs based on the flows,  $V_a^{(n)}$ . If there is no significant change in the flows for two consecutive iterations, then stop. Otherwise proceed to step 2.

Each user has a perceived cost for each link, which is selected at random from a distribution table, whose mean is the actual link cost (Hislop et al., 1991). It has been shown by Sheffi, (1985), that making  $\lambda = 1/n$  produces a solution that converges to Wardrop equilibrium but is less efficient in doing so than the Frank-Wolfe algorithm.

#### 4.5.7 Conclusion

The standard default procedures within the SATURN model are the Frank-Wolfe algorithm used within user equilibrium assignment. These have also been accepted as

the standard for the DTO Traffic Model, which also has multiple user classes. Daganzo (1977) and van Vliet (1976) have argued for using stochastic assignment in relatively uncongested transportation networks while using user equilibrium assignment when dealing with congested transportation networks. Sheffi et al. (1981) have stated that while stochastic models are more realistic in terms of modelling behaviour, they have much larger computational requirements and, the differences between stochastic and user equilibrium flow patterns grow smaller as congestion over the transportation network increases.

The research carried out for this thesis is concentrating on congested transportation networks in the GDA, which means using user equilibrium assignment for the reasons stated in the previous paragraph. Though PARTAN is considerably faster than Frank-Wolfe, this research also needs to look at multiple user classes over congested transportation networks. This excludes the PARTAN algorithm for solving Wardrop equilibrium assignment leaving us with the Frank-Wolfe approach.

## **4.6 Shortest Paths**

### *4.6.1 Introduction*

Due to the fact that this research is concentrating on Wardrop equilibrium assignment for the reasons described in Section 4.5, and because this assignment is to be applied to multiple user classes, this research will concentrate on the adaptation of the Frank-Wolfe algorithm as described in Section 4.5.4 above.

Within the assignment algorithm of the CONTRAM model, Greenwood and Taylor (1993) have shown that the computational time used in the calculation of the shortest paths makes up the dominant part of the overall run time of a single iteration of the dynamic assignment algorithm used. From the results of the profile that can be seen in Table 4.1, it is clear that the shortest path calculations in the SATURN model, using a fixed demand approach, have proven to be just as costly with respect to CPU times. These results have shown that the shortest path algorithm takes up over 80% of the

overall CPU running time for the assignment and simulation program SATALL, see Section 4.4 above.

The shortest paths problem is one of the most well-known network optimisation problems. This problem comes up not just in pure graph theory but in a number of other fields of study including transportation. A number of algorithms to solve the problem have been produced and studied for quite some time; see for example Moore (1957), Bellman (1958), Dijkstra (1959), Ford (1956), Dial (1969), and Pape (1974). Advances have also been made both in theory and evaluation; see for example Ahuja et al. (1990), Goldberg and Radzik (1993), Fredman et al. (1994), and, more recently Cherkassky et al. (1996) and Zhan and Noon (1998).

However, transportation networks have a number of characteristics that distinguish them from other network optimisation problems and pose particular problems in deriving algorithms. Solving for the shortest paths has, according to Hribar et al. (2001), usually been achieved by applying single source labelling algorithms at each source (origin) in the network because of the sparseness of urban road networks (Gallo and Pallottino, 1986; Golden, 1976; Hribar et al., 1995; Pape, 1974).

van Vliet (1978) has described a number of these characteristics and the most widely used algorithms for solving the shortest paths, or tree building, problem on transportation networks and the following Section 4.6.2 describing tree building incorporates aspects of this paper.

#### *4.6.2 Tree building on transportation networks*

There are a number of algorithms that have been developed to solve the shortest path problem, which are relevant to the transport modelling area. van Vliet (1978) has described the most widely used algorithms for solving this problem, the two main algorithms which have been produced are due to Moore (1957), a label-correcting algorithm, and due to Dijkstra (1959), a label-setting algorithm; see Section 4.6.3 for more detail on the use of labelling in shortest path algorithms.

The characteristics of road networks that have consequences for shortest path algorithms can be summarised as follows (van Vliet, 1978):

- Link lengths between nodes, i.e., arcs, are always positive;
- A path from an origin to a destination must use real nodes, i.e., not nodes at centroids; see Section 4.2.3 for more on centroids;
- Some links may exist in only one direction while two-way links may have different lengths;
- Certain turns or movements may be either banned or penalised to discourage their use;
- The ratio of links to nodes is relatively small, being approximately three, and link lengths are usually represented by integers and are short, e.g. less than a hundred; and,
- Networks can be very large with up to 5000 nodes and 15,000 links.

The method for building a minimum path tree may be described using the following general notation:

- The length of a link between two nodes, A and B, in the network is  $L_{A,B}$ ;
- The length of the chosen path is the sum of the link lengths in the path;
- The minimum distance from the origin, O, to the node A is  $L_A$ ; and,
- The back-node of A is  $B_A$  and the link  $(B_A, A)$  is a part of the shortest path.

*Set-up:*

- Set all  $L_A = \infty$ ,  $L_O = 0$  and all  $B_A$  equal to some default value;

- Set up a loose-end table  $T$  to contain nodes,  $T_i$ , already reached by the algorithm but not fully analysed as back-nodes for further nodes. Entries into this table represent the outer ends of the tree as branches grow to connect all of the nodes together; and,
- All entries are set to zero,  $T_i = 0$ .

**Method:**

Starting with the origin, 'O', equal to the 'current node', A;

- (1) Look at each link (A, B) from the 'current node', A, in turn and if  $L_A + L_{(A,B)} < L_B$  then set  $L_B = L_A + L_{(A,B)}$ , set  $B_B = A$ , and add B to the loose-end table, T, provided that B is not a centroid if routing through centroids is prohibited;
- (2) Remove A from T and if the table is empty, stop. Otherwise continue to step (3); and,
- (3) Select another node from the table and return to step (1) with it as the 'current node' A.

Once this process has come to an end then  $L_A$  contains the set of minimum lengths, measured using costs, distances or by other means, from the origin, O, to each node or centroid, A. The back-node,  $B_A$ , then contains all of the data necessary to retrace all of the shortest paths.

#### 4.6.3 Shortest path algorithms for transportation networks

A number of classical algorithms exist which solve the shortest paths problem within the field of transportation. These include the Bellman-Ford-Moore algorithm (Bellman, 1958; Ford and Fulkerson, 1962; Moore, 1958), algorithms due to Moore (1959), Dijkstra (1959), Pallottino (1984) and the D'Esopo-Pape algorithm, as described and tested by Pape (1974) while being credited to D'Esopo by Pollock and Wiebenson (1960). A number of reviews exist on shortest path algorithms in general; for example see, amongst others, Deo and Pang (1984), Gallo and Pallottino (1986), Bertsekas

(1991b), Ahuja et al. (1993) and Cherkassky et al. (1996), and within transportation models see, van Vliet (1978) and Gallo and Pallottino (1984). It is recognised that these algorithms have all developed using different strategies for selecting labelled nodes to be scanned in order to solve the single-source shortest paths problem on transportation networks and other sparse networks (Hribar et al., 2001; Goldberg and Radzik, 1993; Gallo and Pallottino, 1986; Golden, 1976).

van Vliet (1977) found that the set of shortest paths between different zones in a transportation network, commonly referred to as 'building trees', can be a time-consuming problem on a computer, particularly for large networks. The majority of tree building algorithms in general use in transport modelling are of a label-correcting nature, e.g., Moore's algorithm (1957), though one label-setting algorithm is quite prominent also, that due to Dijkstra (1959). The main advantage of Dijkstra over Moore is that provided there are no negative links, it is a once-through method, i.e., building the tree from the source each node and therefore each link is only visited once (van Vliet, 1977). With Moore, a shorter path may be discovered at a later stage than when a node is introduced. This means re-examination of links is necessary. A third algorithm used for this type of modelling but not as common as the other two is called the D'Esopo algorithm (Pollock and Wiebenson, 1960). Its advantage is that it corrects errors as soon as they are detected without allowing them to progress further. van Vliet (1977) found that D'Esopo was the fastest of the three on all but two of the nine networks he tested and the relative efficiency is sensitive to network size and in the case of Dijkstra, link length. D'Esopo can reduce central processing unit (CPU) times by more than 50% relative to standard algorithms in use (van Vliet, 1977). The D'Esopo algorithm is used by the SATURN model, one of the most commonly used transportation network models in the UK and Europe.

Labelling algorithms consist of iteratively updating node labels, where the label represents the shortest path, be that distance or a different measure of cost, at the end of the algorithm. A general label-correcting method has been described by Goldberg and Radzik (1993) for solving shortest path problems. The simplest method of keeping the set of nodes eligible for scanning is to use a list. A queue is a list where additions to the



queue are allowed only at the end of the queue and deletions only at the top (Mondou et al., 1991).

Shortest path algorithms produce trees beginning at sources, or origins, on the network. Each node, 'i', on the network has a label,  $dl(i)$ , which represents the smallest known distance from the source to the node, 'i'. Each nodes label on the network is updated iteratively until, at the end of the algorithm, each label represents the shortest distance from the source node to the node, 'i'. The list, as described in the previous paragraph, contains nodes whose labels are not necessarily the shortest path distances. During each iteration a node is removed from the list and its adjacent nodes' labels are updated. The method of removal of these nodes determines whether the algorithm is label-setting or label-correcting.

The basic label-setting algorithm is due to Dijkstra (1959). In label-setting algorithms the node with the smallest distance label is removed from the list during each iteration. If the costs of travelling from node 'i' to node 'j', i.e., travelling along the arc 'ij', are non-negative then the removed node will never enter the list again. Label-correcting algorithms, on the other hand, do not necessarily remove the node whose label is the minimum. Hence, a node that has been removed may re-enter the list at a later time. Due to the fact that the labels not being set, the distance labels are only correct when, for every arc, 'ij', the distance label for the node 'j' is less than or equal to the distance label  $dl(i)$  plus the cost of travelling along the arc,  $C_{ij}$ . Label-correcting algorithms have been shown to have better experimental performance than label-setting algorithms for sparse networks (Hribar et al., 2001, 1995; Gallo and Pallottino, 1986; Golden, 1976; Pape, 1974).

To examine the performance difference between label-correcting and label-setting two of the main algorithms in the field of transportation for solving the shortest path problem, those due to Moore and Dijkstra, are compared. The main difference between these algorithms lies in the procedure for selecting a labelled node from the loose-end table. Moore's algorithm is a label-correcting algorithm, and Dijkstra's algorithm is a label-setting algorithm. Moore selects the top entry on the loose-end table, which is the

oldest entry in the table. On the other hand, Dijkstra selects the node that is nearest to the origin, which requires additional computation but it does ensure that each node is examined only once. It is well known, according to Ortuzar and Willumsen (1994), that Dijkstra's algorithm is superior to Moore's, especially for large networks, although it is more difficult to program. The D'Esopo variation, however, has been found by van Vliet (1977, 1978) to be both more efficient and faster over a range of sparse transportation network sizes and configurations.

The order in which the nodes are chosen from the list affects the number of removals needed to complete the algorithm, and hence the algorithm's performance. The label-correcting Bellman-Ford-Moore algorithm was the first to maintain the set of labelled nodes in a first-in first-out (FIFO) queue where the next node to be scanned is removed from the top of the queue while the node that becomes labelled is added to the end of the queue. Goldberg and Radzik (1993) put forward a heuristic improvement of the Bellman-Ford-Moore algorithm but recognise that in practice, the D'Esopo-Pape algorithm as described by Pape (1974) and the Pallottino algorithm, both also label-correcting algorithms, perform better than the Bellman-Ford-Moore algorithm.

The Pallottino algorithm is a variant of the Bellman-Ford-Moore algorithm and uses two FIFO queues to store the list; one has a low priority and the other a high priority. Each labelled node appears on one of the queues. The next node to be scanned is removed from the top of the high priority queue or from the top of the low priority queue if the high priority queue is empty. If a node becomes labelled for the first time it is added to the end of the low priority queue and to the end of the high priority queue otherwise (Pallottino, 1984; Golberg and Radzik, 1993; Pallottino and Scutella, 1997). This method is described as a variant of the deque (double-ended queue) list, which was initially developed by D'Esopo and used in the D'Esopo-Pape algorithm as described in the next paragraph. The idea of using this method is that nodes that have already been seen and updated at least once, should be updated again before nodes that have yet to be seen (Hribar et al., 2001).

The D'Esopo-Pape algorithm was identified as performing very well for large as well as small networks by van Vliet (1977). The D'Esopo-Pape algorithm is an extension to Moore's algorithm in that it uses a two-ended loose-end table, also known as a deque list for which additions and deletions are possible at either end, so that a node is entered at one of the ends depending on its status. In this way any node that has already been re-examined but comes up again as the algorithm progresses is immediately examined again because it will have been stored at the top of the loose-end table; see van Vliet (1977) and Pape (1974) for more detail. The status can have three values corresponding to the following situations:

- If the node B has not previously been reached before by the algorithm then it is entered at the bottom of the table;
- If it is on the loose-end table already then no further entry is made to the table; and,
- If it has been entered to the loose-end table, analysed and removed from the table then it is entered at the top of the table so that it can be re-examined immediately.

As shown by Van Vliet (1977), the D'Esopo-Pape algorithm can reduce CPU times by up to 50% when compared to Moore's algorithm. Furthermore, from a similar study made by van Vliet (1978) of these algorithms, over a wide range of network sizes and configurations, the D'Esopo modification to Moore's algorithm gives minimum CPU times when compared with those times obtained from the best implementations of Dijkstra's algorithm. The best implementations of Dijkstra's algorithm used a box-sort in 1977. This is demonstrated in Table 4.2 where the D'Esopo algorithm was found to be fastest in all but two of nine networks tested in van Vliet's (1977, 1978) analysis.

However, since then it has been found by Cherkassky et al. (1996) that for problems with positive arc lengths, Dijkstra's algorithm implemented with a 'double-bucket' data structure is robust, and appropriate implementations are usually competitive when compared to other algorithms. A 'double-bucket' data structure is one where origins

that have been assessed by the algorithm are put in two lists instead of one depending on their status. However, this study does not draw comparisons with the D'Esopo modification to Moore's algorithm for transportation networks nor does it use real road networks for evaluation as van Vliet (1978) does.

**Table 4.2: Table of network dimensions and CPU times (Van Vliet, 1977, 1978).**

Network	Dimensions				CPU times (secs)		
	Nodes	Links	Cent.s	Mean Link Length	Dijkstra a Box- Sort	D'Esopo	Moore
1	256	657	57	(a) 14.5	2.05	1.7	2.1
				(b) 24.9	2.3	1.7	2.1
				(c) 35.6	2.4	1.7	2.1
2	567	1684	81	(a) 10.2	6.0	6.4	9.8
				(b) 26.6	6.6	6.0	9.7
				(c) 73.9	8.0	6.4	9.8
3	1811	4923	95	(a) 10.2	22	27	67
				(b) 38.8	26	25	57
				(c) 210.5	48	31	69

Zhan and Noon (1998) makes reference to Cherkassky et al. (1996) work on Dijkstra's algorithm as a starting point for their work but they use real road networks as opposed to randomly generated networks. It is concluded that the best implementations for solving the single-source 'one-to-all' shortest path problem are those due to Pape (1974) and Pallottino (1984). However, the algorithm attributed to Pape (1974) by Zhan and Noon (1998) is described by Pape (1974) as the D'Esopo variation to Moore's algorithm. A number of different variations of Dijkstra's algorithm are also recommended for different network sizes and configurations; however it appears that no one implementation of Dijkstra's algorithm performs well across both large and small networks.

The D'Esopo-Pape algorithm also has the advantage that it is a much simpler algorithm to program (van Vliet, 1977, 1978). It is only slightly more complicated than the

simplest version of Moore and far less complex than the most efficient versions of Dijkstra, which are only more efficient in very specific circumstances. For more detail on the practical efficiencies of the main algorithmic approaches, see Gallo and Pallottino (1988), Hung and Divoky (1988), Mondou et al. (1991) and Cherkassky et al. (1996).

#### 4.6.4 Conclusion

The shortest path problem is the computationally intensive section of the Wardrop equilibrium approach to assignment and hence, the Frank-Wolfe algorithm. Therefore, when dealing with the shortest paths problem in the SATURN model this research will concentrate on reprogramming the D'Esopo-Pape algorithm in parallel for a number of reasons:

- (1) In the case of transportation models, networks generally have nonnegative arc costs and structured sparse graphs. Therefore, the current deque shortest path algorithm, i.e., the D'Esopo-Pape algorithm, seems to be the best choice of algorithm for transportation applications, excluding specific circumstances where specialised algorithms need to be used (Gallo and Pallottino, 1984; Pallottino and Scutella, 1997);
- (2) Its robustness in being able to solve the shortest paths problem for different network sizes and configurations more efficiently than both Moore, which is only faster for very small networks (less than 75 nodes) and Dijkstra (box-sort), which is only faster for large networks with short nonnegative link lengths (Van Vliet, 1978);
- (3) The D'Esopo-Pape algorithm may be used to prevent paths being routed through centroid nodes; see characteristics of transportation networks in Section 4.6.2. If the status of all of the centroid nodes is set to zero then the algorithm will never enter them in the loose-end table. This means that exit links from centroid nodes, which represent either origins or destinations, are effectively excluded from the tree (van Vliet, 1977). This also means that a computationally

expensive check to ensure that through routing of centroid nodes does not occur is avoided; and,

- (4) Its programming simplicity when compared to that of any of the most efficient variations of Dijkstra's algorithm.

#### ***4.7 Review of parallel programming in the area of traffic assignment and shortest path calculation***

##### *4.7.1 Introduction*

The field of transportation has, over the past half century, pushed against the barriers of computational performance. Traffic equilibrium applications are large, computationally intense programs. However, while the development of single processor speed and power has been impressive, it is insufficient in meeting the computational demands of traffic problems (Hribar et al., 2001). Parallel computing offers an increase in computing power over and above what can be obtained sequentially and is, therefore, a highly relevant field for transportation (Allsop, 1997).

A larger burden has been put on computing power as applications have become more complex and the demand for user friendliness, better graphics and faster evaluation of alternative scenarios has increased. Single processor computing power and processor speeds have steadily increased over the last 50 years. However, a limit is rapidly being approached that is determined by the speed of light. The overall speed of a computer system is not just determined by processor speed however, but by the speed of its slowest components, the proximity of those components to each other and the speed at which data can be transferred between those constituent components. To counteract these limiting factors, electronic components are produced to be as small as industry can achieve, and placed as close together as possible. However, there is also a limit to how close electronic components can be placed together, beyond which they cease to interact properly as individual components.

Parallel computing is a form of computation that offers a solution to these physical limitations for problems that can be broken down into a number of constituent sub-problems, or subtasks. These subtasks are distributed among a number of processors so that they can be operated on simultaneously and solved more efficiently using parallel programming techniques and hardware. Parallelisation also offers a solution to the '*von Neumann bottleneck*', which can occur when all tasks have to be processed in sequence by a single processor; see Riley (1987) for a more detailed description of von Neumann architecture.

However, the application of parallel methodologies and techniques to parallel platforms is considerably more complex than in the sequential world. In the hardware domain one must consider not just the instruction rate of the processors but the number of processors and how they interact. In the software domain one must consider the service demands of multiple processors as well as the structure of the software with respect to the hardware architecture. For instance, it is noted by Chabini et al. (1997) that debugging of parallel programs is an order of magnitude harder than for sequential applications.

This review describes an overview of parallel computing architectures in Section 4.7.2. Section 4.7.3 describes the methods and considerations that need to be understood in order to design and implement parallel programs. The computational methods of data and process decomposition, computational and storage dependence, and, inter-process communication are discussed. Section 4.7.4 describes performance analysis of parallel systems, including a description of speed-up, relative burden, average parallelism and benefit/cost profiles. Parallel processing in traffic assignment and shortest path theory is reviewed in Section 4.7.5 with particular reference to single source one-to-all shortest path parallel algorithms. Chapter 5, Section 5.2, draws conclusions from the parallel computing review in this Section by describing a parallel application methodology and parallelisation strategy for the parallel implementation of the sequential transport network SATURN model, and more specifically the assignment process.

#### 4.7.2 *Parallel computing architecture*

The programmer should try to identify the parallelisation possibilities in an algorithm and then use them to take full advantage of the architecture of the computer system and the development tools available (Chabini et al., 1997).

Hockney et al. (1988) describe two types of supercomputing architecture. The first type of parallel architecture uses pipelining and vector processing. Pipelining is used where separate operations are repeated at each of a number of stages on a stream of data. A stream of data is considered to be a sequence of instructions or data as operated on by a processor. Each operation is carried out by a different processing element, or processing unit. Vector processing is where vectors of numbers are operated on by purpose built hardware, for example the CRAY range of hardware, which uses vector processing and pipelining.

The second type of parallel architecture is the processor array. Flynn (1972) introduced a double dichotomy based on processes using single or multiple streams of instructions and single or multiple streams of data as a classification for parallel systems. Four types of parallel architecture were identified. The first two are Single Instruction, Single Data (SISD) and Multiple Instruction, Single Data (MISD). However, processor arrays fall into the second two groups depending on how they relate their operating instructions to the data being processed. Most commercially available parallel computers fall into these two latter groups:

- (1) **Single Instruction stream / Multiple Data stream (SIMD).** Each processor in the array carries out the same instruction in parallel but on different data streams and processors operate in step with each other; and
- (2) **Multiple Instruction stream / Multiple Data stream (MIMD).** Each processor runs using its own instruction stream and operates on its own data stream, e.g., a transputer array.

Supercomputers such as the CRAY range can cost millions of dollars for onsite installation but can take existing sequential programs written in computer languages



such as C, C++ and FORTRAN and convert the code so that it can run in parallel on the machine with minimal effort on the behalf of the programmer. MIMD machines are far cheaper and offer far greater flexibility in the way problems are distributed, examples being transputers and distributed networks of workstations. However, they are harder to set-up and program in parallel (Hislop et al., 1991). An example of a SIMD machine is the CM-2 connection machine, see Zenios (1989). The majority of parallel systems becoming available belong to the MIMD category however (Allsop, 1997).

There are a number of other hardware characteristics that must be considered when looking at the implementation of an algorithm in parallel. These include the number of processors; fine grained, or massively, parallel systems have in the order of thousands of processors whereas coarse grained parallel systems would usually have a couple of dozen processors. The grain, referred to above, is a description of the size of the data packet that a processor can handle in Random Access Memory (RAM). Fine grained processors have perhaps 16-20 Kbytes of RAM whereas coarse grained processors may have gigabytes of RAM available. Synchronisation refers to the presence of a global clock in the computer system. SIMD machines are generally synchronous whereas MIMD machines are usually asynchronous. Finally, communication may take place among processors using two systems:

- (1) In shared memory machines processors may write/read to a common memory. Shared memory machines have two types: tightly coupled and loosely coupled. In the tightly coupled format one memory block exists and is accessed by all of the processors. In the loosely coupled format each processor has a block of memory that can be fetched by all of the other processors; and,
- (2) Processors may exchange messages using, for example, the MPI (Message Passing Interface) interface to communicate between processors. Message passing systems use an interconnection network over which they exchange messages along links. This network has a topology describing the outlay of the network. This topology must be chosen with the communication requirements

of the parallel algorithm in mind. Common topologies include the ring, the tree, the mesh and the hypercube, or three-dimensional mesh.

#### 4.7.3 *Parallel computation: methods and considerations*

A number of issues need to be understood in order to design and implement parallel programs; see *Parallel Computation: Models and Methods* by Akl (1997). The parallel programmer should look for any explicit parallelisation potential in the sequential code. The programmer needs to be aware that when a sequential program executes the order in which its processes execute is known so the result can be determined easily. However, when a parallelised version executes the order of process completion may vary leading to, potentially, an indeterminate result. Also, every parallel process has an associated optimal processor topology that minimises the amount of inter-process communication. Additionally, it should be noted that not all sequential software is suitable for parallelisation and not all methods of parallelisation perform well, mainly due to communication overheads. However, parallelised software should be scalable, i.e., give consistent results across different parallel architectures.

##### 4.7.3.1 **Decomposition**

Parallel program design methodologies depend on the category to which the parallel system belongs. Decomposition is the most common approach for the design of parallel programs on multiprocessor parallel systems. As such, two types of parallelisation are generally used; data decomposition and process decomposition:

- Data decomposition distributes any global data to all processors and each processor executes the same code on its assigned portion of the data; and,
- Process decomposition breaks down the program algorithm itself and each processor executes a different section. This system is likened to an assembly line and may be referred to as '*data flow*' as data is moved amongst various processes.

Using the decomposition technique each problem, or sub-program within the overall application, is divided into subtasks that can be executed simultaneously over a number of processors in parallel. These problems, or tasks, are obtained through a process that can involve any or all of the following criteria as described by Chabini et al. (1997): partition, allocation, load balancing, communication and synchronisation.

- Partition consists of determining the size of each of the subtasks to be assigned to each processor. Depending on the resulting size of each subtask three levels of parallelisation are produced; fine, medium and coarse grained. In fine grained parallelism each subtask can represent an operating instruction, as compared to coarse grained parallelism where each subtask can represent a procedure or loop. In order to do this the other criteria must also be considered;
- Allocation looks at assigning each subtask to a processor so that priority constraints amongst subtasks are respected;
- Load balancing is tightly interconnected with allocation and partition. The main objective of load balancing is to keep all the processors in use, as far as is practicable, and equally loaded in terms of the execution times of the processes allocated to them. Fine grained subtasks are useful for load balancing but can result in an increase in communication overheads. Coarse grained subtasks can result in long idle times for processors that have finished their subtask and are waiting for other processors to finish so that the algorithm can continue in a synchronised manner;
- Communications are necessary between processors at some stage during any parallel implementation of this kind. This communication detracts from the overall running efficiency of a parallel program by introducing run time latencies as processors must wait for information to be sent and received to and from other processors in order to themselves proceed with their subtasks. Therefore, to optimise any parallel application it is necessary to minimise the communication costs. To do this it is imperative to take the topology of the

interconnection network of the computing hardware into account, i.e., shared memory machines or distributed memory machines:

- **Shared memory** machines have a limited number of processors that all use the same memory bank; and,
  - **Distributed memory** machines use an architecture where each processor has its own memory and each processor is connected directly with all of the other processors in the system.
- Finally, synchronisation is dependent on precedence constraints in the overall application. In effect, synchronisation looks at how tasks are ordered in the application, from the initial inputs to the outputs at the other end. A simple example of this is that variables in a program must be defined and assigned before they can be used. Processes may have to wait until previous tasks have been completed before certain variables can be accessed. The process of doing this can have an effect on the performance of the parallel implementation.

An example of decomposition can be seen in the master/slave parallel model. Here, a master processor makes independent computations into work packets using the method described above and sends them out to a pool of slave processors for computation. Each slave processor is assigned one of the subtasks, or packets, while the master processor is completely dedicated to the co-ordination of the activities of the slave processors and to their communication requirements. The master/slave paradigm works well when work packets are small and the amount of work to be done, or the time for computation per packet, is large. This is to avoid large communication overheads. However, this approach will eventually be limited by the computational bottleneck inherent in a centralised system, which would arise from traffic levels, or the transportation network, becoming too large (Hislop et al., 1991).

### 4.7.3.2 Computational and storage dependence

Implementation of parallel algorithms without due attention to issues of synchronisation, load balancing, partition, etc., can lead to problems of computational and storage dependency, leading to severe inefficiencies and poor performance of the algorithm. Avoiding computational and storage dependence (Ziliaskopoulos et al., 1997) can be one of the more difficult aspects of parallelising an algorithm. As such, the programmer needs to be aware of what these dependencies are. Computational dependence consists of data dependence and control dependence:

- Data dependence refers to the relationship between the order of statements that use or produce the same data; and,
- Control dependence refers to the situation where the order of the execution statements can not be determined beforehand.

Storage dependence refers to the independence of tasks from each other where the access, usage and storage of variables for one task must not interfere with those of another task.

### 4.7.3.3 Inter-process communication

The implementation of a number of subtasks making up part of a larger overall problem depends on the type of parallel computing architecture being used, i.e., shared or distributed memory using perhaps SIMD or MIMD parallel architecture. For example, in a distributed memory message passing system each subtask is modelled as a process and each process has its own address space. Communication happens between processes operating on their own data along channels between exactly one pair of processes. Each process may have many input or output (I/O) channels. This is known as the communication sequential process model (Chabini et al., 1997) for parallel execution, which was initially developed by Hoare (1978) in the 1970's and is the basis for the message passing paradigm. This model is commonly used in communication libraries such as, for example, the MPI interface (The Message Passing Interface Forum,

1994; Walker, 1994). Two basic functions, called 'send' and 'receive', are generally used to send and receive the information along channels between processes.

The MPI interface has been generally accepted as the message passing interface of preference in parallel computing environments (Hempel and Walker, 1999). The main function of MPI is to communicate data from one process to another much as the TCP/IP mechanism does for lower level networks. While message passing provides the most obvious way of programming a physically distributed memory parallel system, it can also be used on shared memory and sequential computer systems and, as such, can be used as the basis for the development of efficient portable programs on all computer architectures (Hempel and Walker, 1999). By the end of the 1980's many programming interfaces had been developed because each parallel computer hardware company produced their own interface. This lack of portability between different machines led to the development of independent programming interfaces such as PARallel MACroS, or PARMACS, (Calkin et al., 1994; Hempel et al., 1994) and the Parallel Virtual Machine environment, or PVM, (Sunderam, 1990; Geist et al., 1994), which gave portability to programs for different parallel computing systems.

The design of portable programming interfaces was based on the design being sufficiently abstract from the individual parallel system hardware requirements. The resulting design and research undertaken to achieve portable programming interfaces provided the experience for the development of MPI. Also, the involvement of the majority of parallel computer manufacturers ensured that no machine was disadvantaged by the MPI specification. Two MPI specifications have been produced; an annotated reference manual by Snir et al. (1996) for MPI-1 with a revision, also by Snir et al. (1998), and an analogy for the MPI-2 specification by Gropp et al. (1998). However, excluding the positive aspects of the development of portable programming interfaces, the main success of the message passing paradigm is due to its efficiency, scalability for large numbers of processors and portability for many different parallel systems and applications (Hempel and Walker, 1999).

PVM is a system designed for managing and coordinating parallel systems. It was produced by researchers at Emory University, the University of Tennessee, Knoxville, and Oak Ridge National Laboratory as a research development project (Geist et al., 1994; Sunderam, 1990). When the development of the MPI interface began PVM was the most popular message passing system in use. It has been regarded that MPI and PVM have been competing to become the message passing standard since then, however, two of PVM's principle developers, Jack Dongarra and Al Geist, were also key to the development of MPI (Hempel and Walker, 1999).

The design objectives for MPI and PVM differed greatly. PVM was designed for use on networks of workstations and problems to do with interoperability and resource management. As a result, its message passing facilities are not very sophisticated. In contrast, MPI's development focused on message passing and is intended to provide high performance on tightly-coupled homogeneous parallel computing architectures, which it has achieved (Hempel and Walker, 1999). Gropp and Lusk (1997) also provide a comparison between PVM and MPI, finding MPI more credible than other message passing libraries because of the consultative process involved in its creation. Development of message passing aspects of PVM were stopped after Version 3.4 in favour of research into distributed, heterogeneous environments for networks of workstations while some of PVM's major strengths, such as its resource management capabilities, have been incorporated into MPI. Today the majority of parallel computer producers support MPI as their primary message passing interface with other interfaces available for reasons of compatibility with legacy codes (Hempel and Walker, 1999).

Traff (1995) has noted that communication is at least an order of magnitude slower using a distributed memory system as opposed to a shared memory one. Though Hempel and Walker (1999) would, for the most part, disagree with Traff's statement, parallel computation using computer architectures without shared memory, i.e., distributed memory, is subject to a trade-off between keeping each processor as busy as possible and the inter-processor time required to do so. As a result linear increases in performance of parallel implementations are rarely, if ever, obtained.

#### 4.7.4 Performance analysis of parallel systems

##### 4.7.4.1 Introduction

The performance of a parallel computer system, combining a software element in the parallel algorithm and a hardware element in the parallel platform, needs to be assessed. The definition of performance depends on the reason behind the use of parallel programming in the first place. In general a reduction in the computation time of the program is sought. Hence, the improvement in performance is measured by comparing the sequential time to run the program on a single processor against that of the parallel application across a number of processors. This measurement of the improvement in the performance is known as speed-up. Other related measures include relative burden and efficiency.

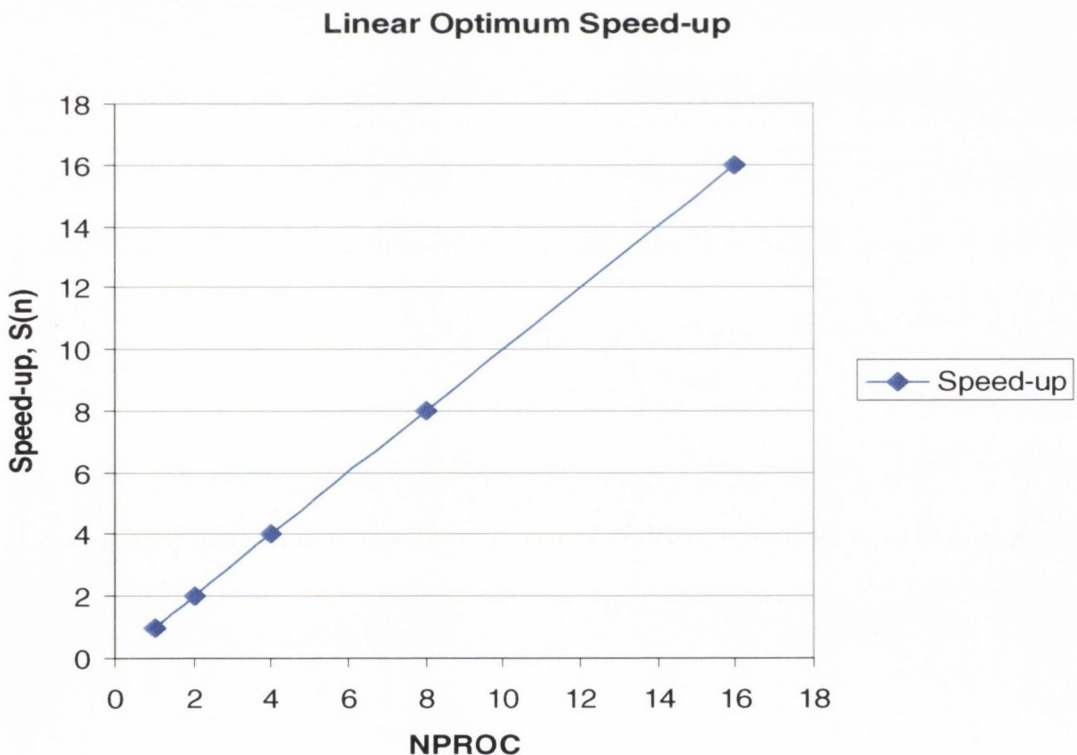


Figure 4.8: The theoretical optimum speed-up for a parallel program.



In an idealised environment the speed-up is linearly related to the number of processors being used to run a parallel application, see Figure 4.9. For example, as the number of processors increases from one to two then the program would run twice as fast and the speed-up would be equal to two. However, due to the sequential nature of the SATURN program and the assignment processes within it, achieving the theoretical optimum, a linear speed-up, is impossible. As the number of processors increases the relative affect of the sequential sections of the parallel model on the performance of the entire model will increase. Therefore, a vastly reduced speed-up is expected. Based on other parallel implementations, for instance the parallel application of the CONTRAM model (Greenwood and Taylor, 1993), a speed-up of approximately 7 times might be expected for the assignment. This research will parallelise the SATURN model, assess the overall performance of this parallel model and also assess the likely affects of the other sequential elements of the model on performance.

#### 4.7.4.2 Speed-up, relative burden and efficiency

Chabini et al. (1997) focuses on the 'execution' of an instance of a problem using the parallel system. Performance measures are reported for a varying size ( $w$ ) of 'execution' and varying the number of processors ( $n$ ) used for that 'execution'. The following parallel performance measures were produced:

- **Speed-up**,  $S(n)$ , is defined as the ratio of the elapsed time, or serial time, when executing a program on a single processor,  $T_s$ , to the execution time when a number of processors,  $n$ , are available,  $T(n)$ . Therefore;

$$S(n) = T_s / T(n)$$

Equation 4.11

This measures the gain in speed between using one processor and 'n' processors to solve a problem. It may also be interpreted as the average number of processors kept busy during the execution of the problem.

- **Relative Burden**,  $B(w,n)$ , (Chabini (1994); Chabini and Florian (1995); Chabini and Gendon (1995)) measures the deviation from the ideal improvement in time between the serial execution and the parallel execution normalised by the serial time. Therefore;

$$B(n) = T(n) / T_s - 1 / n \quad \text{Equation 4.12}$$

The speed-up and relative burden are interrelated with  $S(w,n) = n / (1 + B(w,n)n)$ . Burden graphs are used expressly to compliment speed-up curves.

The idea of using parallelism to speed up the execution of a program has, according to Kuck (1977), existed for more than a century, although the parallel systems that exist today have only become common in recent times. As more processors are assigned to the execution of a problem it is expected that the speed-up will increase. However, it may also be expected that the total idle time will increase due to communication overheads between processors and processes, the software structure, and, contention between different components of the system for shared resources, i.e., two processors trying to access the same memory block at the same time in a share memory parallel computer (Eager et al., 1989). Overheads due to I/O are, in general, not included in the assessment of the performance of the parallel system. The overheads, as described above, are said to be represented by including them in the service demands of the various subtasks. In this way one can assume that they are fixed and do not vary with the number of processors being used nor with the scheduling procedure for the subtasks (Eager et al., 1989).

The efficiency, 'E', is defined by Eager et al. (1989) as the average utilisation of the 'n' allocated processors. Ignoring I/O the efficiency of a single processor system is 1 and hence, the speed-up is equal to 1. The relationship between efficiency and speed-up, described by Eager et al. (1989) as the average processor utilisation, can then be defined by  $E(n) = S(n) / n$ , leading us to the definition of linear speed-up, i.e., where the

efficiency remains at 1 while the number of processors increases beyond two. Linear speed-up is impossible for the same reasons idle time exists, as described above. Studies such as those by Minsky et al. (1971) and Lee et al. (1985) have tried to define the 'typical' speed-up though this is impossible in general terms as the speed-up of each implementation of a parallel algorithm depends on the characteristics of the entire parallel system, which are invariably different for each parallel application.

#### 4.7.4.3 The average parallelism

The average parallelism measure was first introduced by Gurd et al. (1985) in their review of the architecture and performance of the "Manchester Prototype Dataflow Computer". They found that programs with a similar value of average parallelism exhibit virtually identical speed-up curves and that the higher the value, the closer the program got to achieving 100 % utilisation, i.e. an efficiency of 1, on a graph. This seems to indicate that the measure of average parallelism is all that is necessary to ascertain an accurate description of its speed-up curve, regardless of other factors such as the source code language, the time variance of parallelism, etc. Gurd et al. (1985) also concluded that larger applications codes exhibited the same patterns as simpler samples. Eager et al. (1989) define the average parallelism as:

- The average number of processors that remain busy during the execution of a software system given an unlimited number of available processors;
- The speed-up given an unlimited number of available processors, i.e., the maximum possible speed-up;
- The intersection of the hardware limitation with that of the software limitation on speed-up, i.e., two upper bounds on speed-up based on the hardware and software; or finally,
- The ratio of the total service demand (the execution time when a single processor is used or the sum of the service demands of the subtasks) required by the computation to the length of the longest subtask path (where multiple

processors are being utilized), where the length of a path is measured by the sum of the service demands of its subtasks.

The hardware limit on speed-up is given by the number of processors available, 'n', and is only met if all of the processors being used can be kept fully utilized all of the time. The software limitation is derived from the fact that no matter how many processors are used the execution time must be at least as long as the length of a longest path.

Eager et al. (1989) go on to describe the use of the parameter 'f', the fraction of work that is inherently sequential, to tighten both the upper and lower bounds on speed-up. 'f' is defined as the ratio of the service demand of the sequential parts of the computation to the total service demands of the computation where I/O is not included.

#### 4.7.4.4 Benefit verses cost profiles

Profiles that plot benefit against cost are common in many areas. The concept of there being a 'knee' in such a profile is a fundamental one, see Denning (1980). The 'knee' is the point where the benefit per unit cost is maximised. The execution time – efficiency profile is one such cost – benefit profile in parallel systems. Eager et al. (1989) give two motivational stand points for the use of this graph in parallel systems:

- (1) Efficiency is viewed as an indication of benefit (as efficiency goes up so does benefit) and execution time as an indication of cost (as execution time goes up so does cost). The system objective that can be taken from this is to achieve efficient usage of each processor while taking into account the cost to users in the form of increased execution times; and,
- (2) Execution time is taken as an indication of benefit (lower the execution time the higher the benefit) and efficiency is taken as an indication of cost (lower the efficiency the higher the cost). The objective that can be taken from this is to achieve low execution times while taking into account the utilisation cost of low efficiency.

On the execution time – efficiency profile each point represents the execution time versus the efficiency for a certain number of processors. The ‘knee’ occurs where the ratio of efficiency to execution time,  $E(n) / T_n$ , is maximised, where ‘n’ is the number of processors allocated to the computation at this point. Eager et al. (1989) proves that at the ‘knee’ of the execution time – efficiency profile a speed-up,  $S(n)$ , and efficiency,  $E(n)$ , of at least 50% of their maximum is guaranteed. All the processors are being utilised at least 50% at this point but if one more processor is added then it would be utilised no more than 50%. It is further shown that the location of the ‘knee’, i.e., the number of processors used at this location, is well approximated by the average parallelism, where the same guarantees regarding speed-up and efficiency are identical to those at the ‘knee’.

#### 4.7.5 *Parallel processing in traffic assignment and shortest path theory*

##### **4.7.5.1 Assignment**

Greenwood and Taylor (1993) found that parallelisation of the dynamic CONTRAM traffic assignment model offers substantially increased performance to model larger transportation networks and to model a larger range of transport scenarios. The principle results from this parallelisation found that approximately a 7-fold speed-up could be achieved with large networks benefiting more from large parallel systems (Greenwood and Taylor, 1993).

It was found that most benefit came from assigning vehicles to their routes in parallel because it involves physically independent or simultaneous processes. Significant changes to the methodology of the model were required to ensure convergence and computational efficiency though the essential characteristics of the CONTRAM model remained the same (Greenwood and Taylor, 1993). The parallel approach used was a “farmer/worker” paradigm, otherwise known as the master/slave paradigm described previously. However, Hislop et al.’s (1991) approach where the network is dissected into zones and traffic is assigned to the zones in parallel was discarded because of the difficulties with assigning ‘packets’ to cross-boundary routes. It was decided that a simpler data decomposition strategy to assign all ‘packets’ in parallel was more

appropriate. This resulted in a speed-up of nine times using 16 T805 Transputers for a network of 1001 links.

Chabini et al. (1997) published a paper entitled 'Parallel and distributed computation of shortest routes and network equilibrium models' in 1997. This paper presents parallel computing implementations of the linear approximation method for solving the fixed demand network equilibrium problem. The paper also contains parallel computing implementations of a shortest paths algorithm due to Dijkstra (1959).

The linear approximation method used to solve the fixed demand network equilibrium problem by Chabini et al. (1997) is an adaptation of the Frank-Wolfe algorithm (Frank and Wolfe, 1956; Arezki and van Vliet, 1990). In the sequential implementation of this algorithm a variant of Dijkstra's (1959) algorithm is used to calculate the shortest routes for all origin-destination pairs. The following paragraph describes, briefly, the methodology used to parallelise Dijkstra's algorithm, as developed by Chabini et al., (1997).

Chabini et al. (1997) uses a master/slave decomposition model, as described in Section 4.7.3 above, to parallelise Dijkstra's algorithm. The PVM environment (Geist et al., 1994; Sunderam, 1990), which is predominantly used for message passing between processors on distributed memory systems and usually for networks of workstations, is used to manage inter-processor communication and the parallel process synchronisation. PVM is also used for portability as it supports low-level networks of heterogeneous workstations when implementing parallel process based algorithms. In order to parallelise the algorithm the calculation of the shortest paths is considered to be a loop over the origins. Dividing the problem into a number of subtasks that can then run on separate processors allows this loop to be partitioned across the network of processors. Each subtask then corresponds to the computation of the shortest paths for a subset of the total number of origins. Once all of the subtasks are completed the output information is transferred back to the master processor to continue with the application.

Chabini et al.'s results indicated that parallel computing offers significant advantages for the parallel computation of shortest paths and for solving the fixed demand network

equilibrium problem on the computing platforms and environments that were available at the time.

Nagel and Rickert (2001) make use of the master/slave approach using decomposition for a parallel implementation of the TRANSIMS micro-simulation. Reference is also made to Chabini (1998) where domain decomposition is used to partition the network graph into domains of approximately the same size for a discrete dynamic shortest path problem.

Hislop et al. (1991) describe two approaches to parallelising traffic assignment, one based on 'central control' and the other on 'distributed control'. The 'central control', or data decomposition, approach is the same as the master/slave paradigm described above, which would work well for the assignment process as each slave processor would only need the global information of origin co-ordinates and traffic volumes on-route to each destination. While this approach may not be the theoretical optimum strategy it would retain the overall framework of the model (Greenwood and Taylor, 1993). The Transport Parallel Computing Centre at University College London has implemented this on a 36 \* T800 transputer array. A relatively good performance was obtained for the size of network used (Hislop et al., 1991). Hislop et al. (1991) also mention using a "geometric parallelism" for the simulation aspect of assignment where the data domain is divided up into regions and mapped onto the processor array so that flows, queues and link costs in each region could be calculated in parallel.

Three proposed methods are given for the distributed approach.

- (1) The first is a NEMIS-like system where a processor is responsible for each node, which is not very economical.
- (2) The second is a regional distribution where processors control different regions of the network. Each processor maintains link costs and flows for its region and a less detailed copy of link costs for the rest of the network, like the buffer and simulation networks in SATURN. Each vehicle would have to be 'moved' as

far as it could go within a discrete time interval instead of being assigned to its entire route per assignment iteration.

- (3) The third method is the same as the second except that instead of a processor having a less detailed copy of the network outside of its region, a distributed minimum cost tree-building algorithm would operate.

Problems with these methods are associated with large communication overheads incurred while transferring network information between processors. The third method could also potentially have problems calculating routes for large networks

#### **4.7.5.2 Single source one-to-all shortest path parallel algorithms**

The majority of international research that has taken place has focused on developing and comparing different parallel implementations for sequential label-setting and label-correcting algorithms (Hribar et al., 2001). A number of parallel algorithms have been proposed for various sequential shortest path algorithms such as those documented by Tseng et al. (1990), Paige and Kruskal (1985) and Chandy and Misra (1982) on machine independent algorithms and those by Habbal et al. (1994) and Dey and Srimani (1989) on machine specific algorithms. The most investigated area being the all-pairs problem, see Habbal et al. (1994), Paige and Kruskal (1985), Chandy and Misra (1982) and Deo et al. (1980). Several attempts have been made to parallelise the single source shortest path problem but with relatively little success (Traff, 1995; Bertsekas and Tsitsiklis, 1989; Paige and Kruskal, 1985; Deo et al., 1980), however, certain adaptations of Moore's algorithm seem promising according to Deo et al. (1980).

By 1991, parallel algorithms for finding the shortest paths within the field of transportation were based on Dijkstra, Moore and D'Esopo, and, Dynamic Programming (Hislop et al., 1991). Bitz and Kung (1988) distributed Dynamic Programming on an iWarp systolic array but found that the method did not perform well on large networks. Mateti and Deo (1982) give examples of distributing Dijkstra's and Moore's algorithms on a shared memory MIMD machine and on a specialised vector addition and comparison supercomputer. Further to this Deo et al. (1983) have



published results showing near linear improvement for the number of processors used for the algorithms distributed on a MIMD shared memory architecture.

Habbal et al. (1994) present an algorithm for a distributed network achieving speed-up but acknowledging that the performance is dependent on the decomposition of the network. Paige and Kruskal (1985) present synchronous parallel versions of Dijkstra's and Ford's algorithms but without any actual implementation or results. Traff (1995) has compared two asynchronous distributed label-setting algorithms based on Dijkstra's algorithm and found that both algorithms initially performed poorly but, when compared to Paige and Kruskal's (1985) synchronous implementation with measures included to reduce the communication costs, performed well. Experimental studies of distributed label-setting algorithms, such as those by Traff (1995) and Adamson and Tick (1991), use a partition of the network among processors so that each processor has its own sub network on which to work. It was observed that label-setting algorithms have little parallelism (Hribar et al., 2001). Mehlhorn et al. (1998) also describe a theoretical approach to parallelising Dijkstra's algorithm by dividing the algorithm into a number of phases to be executed in parallel, a method of process decomposition. The implementation shows good theoretical behaviour and further research is recommended. Bertsekas and Tsitsiklis (1991) have surveyed both synchronous and asynchronous iterative algorithms with mixed convergence results. Traff (1994) also found that for distributed memory systems achieving speed-up for Moore's algorithm is 'apparently' easier than for Dijkstra's.

Two data parallel implementations of Floyd's algorithm are discussed by Narayanan (1992) where the algorithms are compared to each other but there is no measure of the speed-up of either implementation relative to a sequential algorithm. Bertsekas and Polymenakos (1994) have used an auction algorithm developed by Bertsekas (1991), and, Bertsekas and Castanon (1991). The auction algorithm is itself slower than standard sequential single source shortest path algorithms but is reported to be suited to parallelisation and yields speed-up results that are comparable to those of Traff (1995) for a shared memory machine. Other reports have reported on experimental results for '*simulated*' shared memory computers; see Adamson and Tick (1991) and Bertsekas et

al. (1996) as examples, while there appear to be few reports for the more realistic distributed systems; see for example Papaefthymiou and Rodrique (1994). Bertsekas et al. (1996) compares several label-correcting algorithms while Papaefthymiou and Rodrique (1994) implement parallel versions of the label-correcting Bellman-Ford-Moore algorithm and compare it to a sequential implementation of the Bellman-Ford-Moore algorithm developed by Cherkassky et al. (1996).

The main problem with parallelising the single source shortest path problem is that the problem is inherently sequential with a large number of iterations taking short periods of time yielding very small grain sizes. One way to overcome this problem is to perform more computations within every iteration, so as to increase the grain size (Ziliaskopoulos et al., 1997).

Ziliaskopoulos et al. (1997) introduces parallel designs for the dynamic time-dependent least time path algorithms using shared memory and message passing approaches. The parallel algorithms are based on sequential dynamic algorithms introduced by Ziliaskopoulos and Mahmassani (1993, 1996) using discrete time steps and the theory of Bellman's principle of optimality. However, no testing was carried out on large networks. This research is continued through Ziliaskopoulos and Dimitrios (2001) who propose and execute a massively parallel design for the time-dependent least time path algorithm, which is also based on sequential work by Ziliaskopoulos and Mahmassani (1993, 1996), on a PVM system. While speed-up improves with the number of time intervals it is also acknowledged that further research into efficiency, over actual and random networks, is necessary.

A number of papers have been published by Hribar et al. (2001, 1998, 1995) looking at performance aspects of parallel shortest path algorithms and dynamic implementation within the transportation field. Hribar et al. (2001) explore various network decompositions to develop shortest path algorithms and the choice of shortest path algorithm to solve route choice on congested networks but the results are found not to be superior to a simpler decomposition by origin (Florian, 2001). Hribar et al. (1998) examines termination detection implementation, the third aspect of what they have

determined to be the parallel implementation issues of applying simple labelling algorithms to distributed memory machines to solve transportation problems. However, Hribar et al. (2001) focuses on “why algorithms perform the way they do” as opposed to the type of performance analysis described in Section 4.7.4 of this thesis. A distributed network approach, where the network is partitioned into ‘n’ sub networks for ‘n’ processors, is used for solving the shortest path problem using labelling algorithms on distributed memory machines. This takes advantage of the aggregate memory of the distributed system and parallelism can exceed the number of sources. However, there are high communication overheads associated with communicating node labels, and for termination detection (Hribar et al. 2001).

Another approach is to use network replication (Kumar et al., 1991) where the entire network is assigned to each processor and each processor solves the shortest path problem for a subset of sources (origins). There is no inter-processor communication required for this solution. However, a global communication is required to update the replicated network across the processors. The disadvantages are that the parallelism is limited to the number of sources and that the communication overhead increases as the number of processors increase. Hribar et al.’s (2001) approach uses the Single Program, Multiple Data (SPMD) model, where each processor solves the shortest path problem for all sources across its sub network. Multiple processors share boundary nodes while the remaining nodes are known as interior nodes. The program is terminated when all the processor lists are empty. Only 32 sources were used for ease of analysis where transportation networks can have upwards of hundreds or thousands of sources to solve for. Tremblay and Florian (2001) state that computing implementations based on topological subdivisions of the network are not likely to be more efficient than the decomposition by destination approach. Tremblay and Florian (2001) look at temporal, or dynamic, shortest paths quoting Ziliaskopoulos and Mahmassani (1992, 1993, reported in 1997) as the only previous contributors to the area, see above. The parallelisation strategies implemented use a decomposition-by-origin approach with good results found on a shared memory machine.

*4.7.6 Conclusion*

Chapter 5, Section 5.2, draws conclusions from this parallel computing by describing a parallel application methodology and parallelisation strategy for reprogramming the sequential SATURN model, and more specifically the assignment process.

## *Chapter 5*

# *Parallel Implementation of a Transport Network Model*

## 5.1 Introduction

The objective of the research on which this Chapter reports is to assess the SATURN model (van Vliet, 1982), as applied to the Dublin Transportation Network Model (DTNM) with a view to finding the areas within the model that take up the most CPU time, and to reprogram these areas to run in parallel so that higher levels of efficiency can be achieved (Dublin Transportation Initiative, 1995). Reprogramming the most time consuming, or computationally intensive, areas of the sequential model in parallel should allow for a more efficient use of computer time over a variety of network configurations and sizes, and in doing so, result in improved efficiency in examining the DTNM. The overall objectives for the parallel implementation are to:

- Demonstrate that the modified parallel model converges and produces results that are comparable to the sequential original; and to,
- Evaluate the efficiency of the parallel implementation when more than one processor is used, i.e., 2, 4, 8, and 16 processors.

A detailed review of parallel programming in the areas of traffic assignment theory and shortest path calculation has been undertaken in Chapter 4, Section 4.7. Additionally, the initial analysis of the sequential SATURN model is detailed in Section 4.4. Further to this, the computationally intensive area of trip assignment is assessed in Section 4.5, and that of the shortest path algorithm in Section 4.6.

Section 5.2 draws conclusions from Chapter 4 and the parallel computing review in Section 4.7. The parallel application background, methodology and parallelisation strategy for reprogramming the sequential SATURN model, and, more specifically, the assignment process is also described in Section 5.2. Further to this, Section 5.3 provides results and discussion from the implementation of the parallelisation strategy. Finally, conclusions drawn from the results and discussion in Section 5.3, of the parallelisation implementation, are given in Section 5.4.

## 5.2 *Method and strategy for the parallelisation of the sequential SATURN model*

### 5.2.1 *Introduction*

The major objective to be achieved by reprogramming the sequential SATURN model in parallel is to improve the model's performance so as to be able to apply the model to computationally intensive, and hence time intensive, transportation problems that otherwise might not have been assessed.

In order to achieve this objective a parallelisation strategy was produced. The strategy for programming the SATURN model in parallel is to take the most CPU computationally intensive components of the model, which were found to be the calculation of the assignment and shortest path algorithms, as described in Section 4.4, and to parallelise these sections of the sequential model in such a way that the resulting parallel version of the model: converges; is accurate; has a more efficient performance; is scalable from a parallel computing point of view; and, is portable amongst different parallel computing architectures.

Greenwood and Taylor (1993) give a number of steps to help remodel a sequential program to run in parallel on a high performance computing (HPC) system. These steps have been used as a basic guideline for the strategy to parallelise the sequential SATURN model. These steps are to:

- (1) Profile the time spent on each task;
- (2) Study the existing algorithm and analyse the data flow;
- (3) Investigate methods of parallelisation;
- (4) Design and implement the parallelisation software; and to,
- (5) Evaluate the performance compared to the standard sequential version.

So far the completion of step (1) has been demonstrated with the profiling of the sequential SATURN model resulting in a clear determination of the areas of the model that should be researched with respect to parallelisation; see Chapter 4, Section 4.4 for

more detail. Step (2) has been completed through the analysis of the methods and algorithms used within the assignment and shortest path areas of the sequential model; see Chapter 4, Section 4.5 and Section 4.6, respectively. Chapter 4, Section 4.7 addresses step (3) and part of step (4) through a literature review of parallel computing and high performance computers in the areas of traffic assignment and single-source all-to-one shortest path algorithms within the field of transportation. Step (4), the design and implementation of the parallel software, is addressed in this Section with the strategy used to implement the parallel software. Finally, step (5) is completed in Section 5.3, which gives results of the parallel implementation, evaluates its performance and compares the parallel model with the standard sequential version.

Section 5.2.2 describes the background to the application of parallel theory and methodology to the sequential model and briefly mentions the work that was completed before parallel coding techniques could be applied to the code. Section 5.2.3 describes the hardware and operating system used in the parallel application and performance assessment. Section 5.2.4 details the parallelisation strategy and Section 5.2.5 describes the methodology of programming the single-source one-to-all shortest path algorithm in parallel.

### 5.2.2 Background

Initially an understanding of the code of the sequential SATURN model was necessary in order to understand how to go about parallelising it. From Section 4.4 it was found that the nested subroutine used to solve the shortest path problem, '*desopo\_cb*', demands over 80% of the overall CPU running time for the SATALL program. In addition to this, the subroutine, '*loadit\_plus*', (used to load, or assign, the trips onto these shortest paths) which calls the nested shortest path subroutine, '*desopo\_cb*', uses an additional 5% of the total CPU running time. In Section 4.5 an understanding of the assignment algorithms and methodologies is described and in Section 4.6 shortest path algorithms are assessed in some detail.

The sequential SATURN model, and hence, the computationally intensive assignment/simulation SATALL sub-program, is programmed using the Fortran 77



programming language and compiled using the Salford Fortran 77 compiler for 32-bit Intel microprocessor systems (Salford, 1996). Using the raw code obtained from Dr. Dirck van Vliet, of the SATURN model, the SATURN sub-programs, described in Section 4.3, were compiled on an Intel microprocessor machine, a PC.

Once a detailed understanding of the Fortran code of the sequential model had been obtained the code, designed for the Intel microprocessor and the DOS operating system, was transferred, or ported, to a UNIX based operating system with a RISC microprocessor, see Section 5.2.3. The two types of microprocessor are incompatible in the way in which they store their data, i.e., 'little endian' versus 'big endian', see Appendix 2 for more detail. Therefore, the raw code and input files had to be processed to ensure that they would run properly. Once this was achieved the code was again compiled using 'g77', a UNIX based Fortran compiler, until the sequential SATURN program suite was fully executable. Once this had been achieved the sequential model was profiled, see Section 4.4, resulting in the assessment of the most computationally intensive routines, the '*loadit\_plus*' and '*desopo\_cb*' routines, which are the assignment and shortest path algorithms, respectively.

To make sure that every part of the code that must be considered had been looked at carefully, and to gain a detailed understanding of the routines involved in the input/output (I/O), assignment and shortest path algorithms of the model, large portions of the Fortran code, and in particular the two main subroutines of interest, '*loadit\_plus*' and '*desopo\_cb*', were reverse engineered. What this means is that each line of the code, as seen in the routine itself, has been described and explained in a paragraph directly following the relevant line of code, see Appendix 3 for detail of the '*loadit\_plus*' and '*desopo\_cb*' routines. This was seen as the easiest way to ensure a comprehensive understanding of what the variables, parameters and arrays are doing within the two subroutines of the SATALL program.

At this stage different parallelisation techniques and methodologies were considered for the parallelisation of the sequential SATURN model and in particular the assignment

and shortest path areas, see Chapter 4, resulting in the strategy and methodology described in the following Sections.

### 5.2.3 *The IBM SP2 supercomputer*

The HPC system that was used during this research is an IBM 9076 Scalable POWERParallel SP2 supercomputer and is jointly owned by Trinity College Dublin (TCD) and Queens University Belfast (QUB). It is housed at QUB and consists of 48 nodes with varying amounts of resources assigned to each node. A summary of the system parameters for the IBM 9076 model can be seen in Figure 5.1.

The IBM 9076 parallel SP2 System is a distributed memory multi-processor parallel system composed of processor nodes, which are RISC (reduced instruction set computer) System/6000 processors. A microprocessor is a processor embedded on a

<b>9076 SP2</b>	
<b>Clock cycle</b>	15 ns
<b>Theor. peak performance:</b>	
Per Proc. (64-bit)	0.267 Gflop/s
Maximal (64-bit)	34.1 Gflop/s
Memory/node	64–512/2048 MB (see below)
<b>Communication bandwidth:</b>	
Point-to-point	20+ MB/s
Bisectional	25 GB/s
No. of processors	8–128

**Figure 5.1: System parameters of the IBM 9076 SP2 system.**

microchip. RISC microprocessors are designed to perform a smaller number of computer instruction types so that it can operate at a higher speed, i.e., perform more millions of instructions per second (MIPS). An instruction is an order given to a processor by a computer program. At the lowest level, each instruction is a sequence of '0's and '1's that describe a physical operation the computer is to perform, for example to 'ADD' one number to another. Also, depending on the type of instruction, the specification of special storage areas called registers can be included, which may contain data to be used in carrying out the instruction, or the location in computer memory of data. Since each instruction type that a computer must perform requires

additional transistors and circuitry, a larger list, or set, of computer instructions tends to make the microprocessor more complicated and hence, slower in operation.

John Cocke of IBM Research in Yorktown, New York, originated the RISC concept in 1974 by proving that about 20% of the instructions in a computer did 80% of the work. The RISC concept has led to a more thoughtful and efficient design of the microprocessor.

Besides performance improvement, some advantages of RISC and related design improvements are that:

- A new microprocessor can be developed and tested more quickly if one of its aims is to be less complicated;
- The operating system and application programmers who use the microprocessor's instructions will find it easier to develop code with a smaller instruction set;
- The simplicity of the RISC system allows more freedom to choose how to use the space on a microprocessor; and,
- Higher-level language compilers produce more efficient code than formerly because they have always tended to use the smaller set of instructions to be found in a RISC computer.

The RISC microprocessors are held in nodes that are housed in frames and are interconnected by one or several communication networks, such as Ethernet, token ring, or FDDI (Fibre Distributed Data Interface).

- FDDI is a set of ANSI and ISO standards for data transmission on fibre optic lines in a local area network (LAN) that can extend in range up to 200 km (124 miles). The FDDI protocol is based on the token ring protocol. In addition to being large geographically, an FDDI local area network can support thousands of users.

- A token ring network is a LAN in which all computers are connected in a ring or star topology and a token-passing scheme is used in order to prevent the collision of data between two computers that want to send messages at the same time. The token ring protocol is the second most widely used protocol on LANs, after Ethernet. The IBM Token Ring protocol led to a standard version specified as IEEE 802.5.
- Ethernet is the most widely installed LAN technology. It has been specified as a standard in IEEE 802.3 and was originally developed by Xerox. An Ethernet LAN typically uses coaxial cable or special grades of twisted-pairs wires between processors.

The SP2 is accessed and managed through a 'front-end' control workstation that uses an IBM RISC System/6000 processor running the Parallel System Support Programs (PSSP), which include software functions to monitor the SP2 from a single point of control. The control workstation also monitors system failures. Failing nodes can be taken off line and exchanged without interrupting service.

The operating system used on the IBM 9076 is AIX, IBM's Unix variant. The clock cycle for processors is 15ns giving a peak performance of 266Mflop/s per node. Configurations are housed in frames, or columns, that contain between 8 and 16 processors, depending on the type of node being used. There are two different types of node: thin nodes and wide nodes. There are a number of differences between these two types of node. Wide nodes have twice as many I/O (input/output) channel slots as thin nodes, with 8 instead of 4. They also have the option of having far more memory with a maximum of 2GB as compared to the maximum for thin nodes of 512MB. More important in terms of performance is the fact that the data cache of a wide node is four times larger than that of a thin node (256 KB instead of 64 KB) and that the memory bus is two times wider than that of a thin node, having 8 instead of 4 words per cycle. These two differences have led to a performance gain of a factor of 1.5 for wide nodes over thin nodes. More details about the parallel computing system hardware of the SP2

can be found on the TCHPC (Trinity College Dublin High Performance Computing) web site accessible from TCD's homepage (<http://www.tcd.ie>).

#### 5.2.4 *Parallel strategy*

Any parallel strategy for the parallelisation of the sequential SATURN model shall, in modifying the relevant algorithms and code, ensure that the parallel model can be used on a number of different systems with a minimum of difficulty. Initially the strategy must be to try to identify any parallel potential that the algorithms of the sequential model might have and then use them to take full advantage of the HPC system architecture and the development tools available. Although this research will be applying the new parallel model solution on an IBM RS6000 POWERParallel SP2 supercomputer, see Section 5.2.2, the final solution will be compatible with the majority of parallel HPC systems. The IBM SP2 has 48 Power2SC processors, although only 16 were utilised in this research due to administrative restrictions resulting from the large workload on the machine. In order to ensure this portability amongst HPC systems the solution has been designed to be compatible with both distributed memory and shared memory systems. Any solution designed in this way can be applied to a number of different parallel systems, including heterogeneous networks of workstations.

The most important aspect to consider in the design of the parallel implementation, in order to achieve complete portability between distributed memory and shared memory parallel systems, is how inter-process communication is to be achieved. Inter-process communication has been discussed in some detail in Section 4.7.3, along with related topics such as computational and storage dependencies. It was found that the MPI interface is the most suitable parallel programming message passing interface for high performance on tightly-coupled homogeneous parallel computing architectures (Hempel and Walker, 1999). While the development of the PVM environment has branched off into research of distributed heterogeneous environments, MPI has incorporated a number of positive developments from within PVM, such as resource management capabilities, and is now the primary message passing interface available. Gropp and Lusk (1997) provide a comparison between PVM and MPI finding MPI more credible

than other message passing libraries because of the consultative process involved in its creation. Today the majority of parallel computer producers support MPI as their primary message passing interface with other interfaces available for reasons of compatibility with legacy codes (Hempel and Walker, 1999).

For these reasons the MPI interface is used for message passing between processes, managing processor co-ordination, and, for resource management in the parallel implementation; see Appendix 4 for a description and sample of some of the MPI calls used to program the sequential model in parallel. The MPI interface has a number of more specific advantages that are relevant to any parallel implementation:

- The MPI interface gives a high performance of message passing for communication between processes, therefore reducing inefficiencies associated with communication overheads;
- It is flexible, allowing multi-tasking to be incorporated into the implementation of the parallel solution so that the parallel model can, if the system user requires it, run in the background instead of monopolising computer resources; and,
- It is portable, see Section 4.7.3, over different parallel systems, requiring little adjustment.

The closest available research on parallel implementation for the assignment area, with respect to the SATURN traffic assignment and simulation model, is due to Chabini et al. (1997) where a master-slave paradigm is implemented. Two types of implementation were carried out: one using a distributed memory platform and message passing and the other using a shared memory platform and threads. The shared memory implementation has limited scalability due to computational dependence on a single shared memory block while the distributed implementation uses PVM as the message passing interface. See Kumar and Gupta (1994) for an analysis of the scalability of parallel algorithms and architectures.

After researching a number of potential strategies for the parallel implementation of the sequential SATURN model a Single Program, Multiple Data (SPMD) paradigm was chosen for this research, which involves a data decomposition around the outside of the shortest path algorithm. The SPMD paradigm is an approach where each processor runs the same program but acts on a different set of data. This means that the process of data decomposition has been chosen over process decomposition, see Section 4.7.3, because of the linear nature of the shortest path algorithm and trip assignment processes, and, because of the nature of the input data, i.e., the fact that the data can be split into smaller packets of input and acted upon independently of the other data packets. See the following Section 5.2.4 for more detail. The SPMD approach will be more efficient and scale better over multiple processors.

Each data packet is made up of a number of origin-destinations, which are then acted upon the processor. Technically, at least as many processors could be used as there are origin-destinations on the network. In this theoretical case, each processor would be given a single origin and would find the shortest path from this origin to every other destination and load traffic onto these routes as appropriate.

In addition, due to the sequential nature of the shortest path algorithm and the assignment process, it was found that it was not feasible to attempt a parallelisation of the algorithms themselves.

#### *5.2.5 Programming the single-source one-to-all shortest path algorithm in parallel*

The shortest path problem has been shown to take up over 80% of the overall CPU time for SATALL, see Section 4.4. For every source (origin) in the network a shortest path is found to every possible destination, so the same type of computation is calculated to achieve a shortest path for every Origin-Destination (O-D) pair. In programming the shortest path problem in parallel a simple decomposition by origin is used where the total number of origins is divided into a number of subtasks in such a way that each subtask can be issued to a separate processor for analysis. Each subtask then corresponds to the computation of a subset of the total number of origins and has the same computer code as all the other subtasks. The label-correcting single-source

D'Esopo modification to Moore's algorithm is being used as the shortest path algorithm of choice. Though not the fastest algorithm for specific network sizes and configurations the algorithm performs very well over a wide range of transportation applications unlike other well known algorithms such as Dijkstra's label-setting algorithm, which needs to be invoked in different ways to achieve good performance depending on the specific circumstances, see Section 4.6 for more detail. Chabini et al. (1997) have found that a slow sequential algorithm can lead to very good speed-ups when the serial version of the algorithm is coded and the computation of shortest paths is shared for subsets of origins among processors.

From a hardware perspective fine-grained, or massively, parallel systems have in the order of thousands of processors whereas coarse-grained parallel systems would usually have a couple of dozen processors. This definition makes the IBM SP2 HPC system a coarse-grained parallel system. The grain, referred to above, is a description of the size of the data packet that a processor can handle in Random Access Memory (RAM). Fine-grained processors have perhaps 16-20 Kbytes of RAM whereas coarse-grained processors may have gigabytes of RAM available. For instance, the wide nodes of the IBM SP2 can have as much as two gigabytes of RAM available.

From a software perspective, the use of partition in decomposition consists of determining the size of each of the subtasks to be assigned to each processor. Depending on the resulting size of each subtask three levels of parallelisation are produced: fine, medium and coarse-grained. In fine-grained parallelism each subtask can represent an operating instruction, as compared to coarse-grained parallelism where each subtask can represent a procedure or loop. In order to apply this decomposition methodology in this research other criteria must also be considered, see Section 4.7.3. For instance, allocation looks at assigning each subtask to a processor so that priority constraints amongst subtasks are respected. Load balancing is tightly interconnected with allocation and partition: the main objective of load balancing is to keep all the processors in use, as far as is practicable, and equally loaded in terms of the execution times of the processes allocated to them. Fine-grained subtasks are useful for load balancing but can result in an increase in communication overheads. Coarse-grained



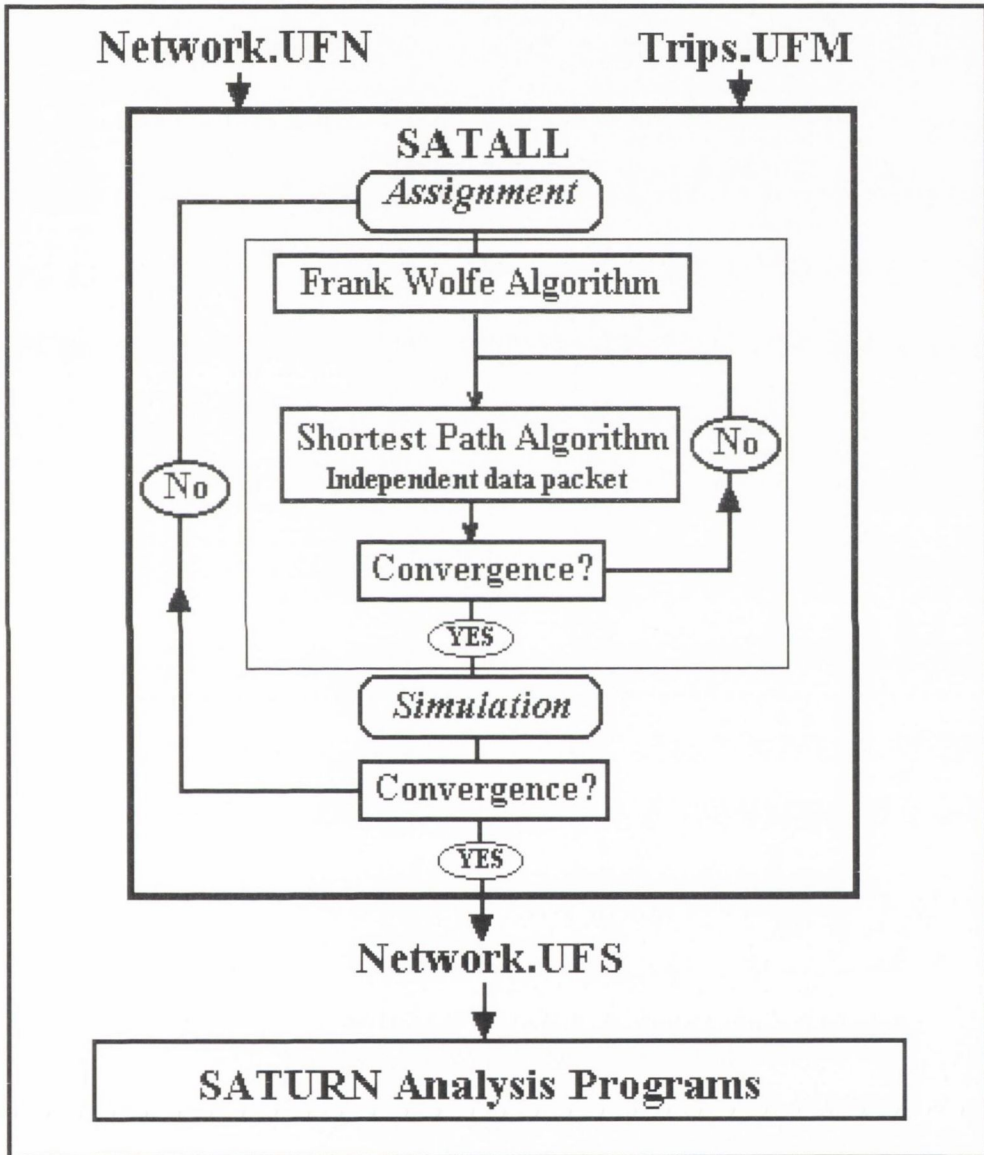
subtasks can result in long idle times for processors that have finished their subtask and are waiting for other processors to finish so that the algorithm can continue in a synchronised manner.

The main problem with parallelising the single-source one-to-all shortest path problem is that the problem is inherently sequential, with a large number of iterations taking short periods of time, yielding very small grain sizes. One way to overcome this problem is to perform more computations within each subtask, so as to increase the grain size (Ziliaskopoulos et al., 1997).

Therefore, in order to parallelise the sequential SATURN model it was decided to create a set of '*coarse-grained*' subtasks based on the fact that the shortest path routine is called so frequently and is so computationally intensive. In addition to this, each shortest path computation is independent of every other shortest path computation, which means that shortest path calculations for different origin-destination (O-D) pairs can be executed simultaneously. Finally, it has been established that the calculation of the shortest paths is an independent process on each processor because it can be calculated on different input data simultaneously. For the O-D pairs sub-tasked to each processor, the trips associated with the O-D pairs can also be assigned to the shortest paths in parallel. This is possible because the other processors are also performing the same task simultaneously, in parallel. Consequently, this is achieved without the need for inter-processor communication.

Figure 5.2 shows a simplified overview of the data flow for a single processor involved in a parallel application of the SATALL program, acting on a coarse-grained data packet. Each processor involved in the parallel application of the model goes through this process with its own 'Independent data packet'. The inter-process communication is not described on Figure 5.2 below but is shown on Figure 5.3. In Figure 5.2 the Network.UFN and Trips.UFM files are input to the SATALL program. SATALL is made up of the assignment and simulation stages. Within the assignment stage the Frank-Wolfe algorithm performs the assignment process until either a convergence parameter or equilibrium is achieved. When this occurs the program progresses to the

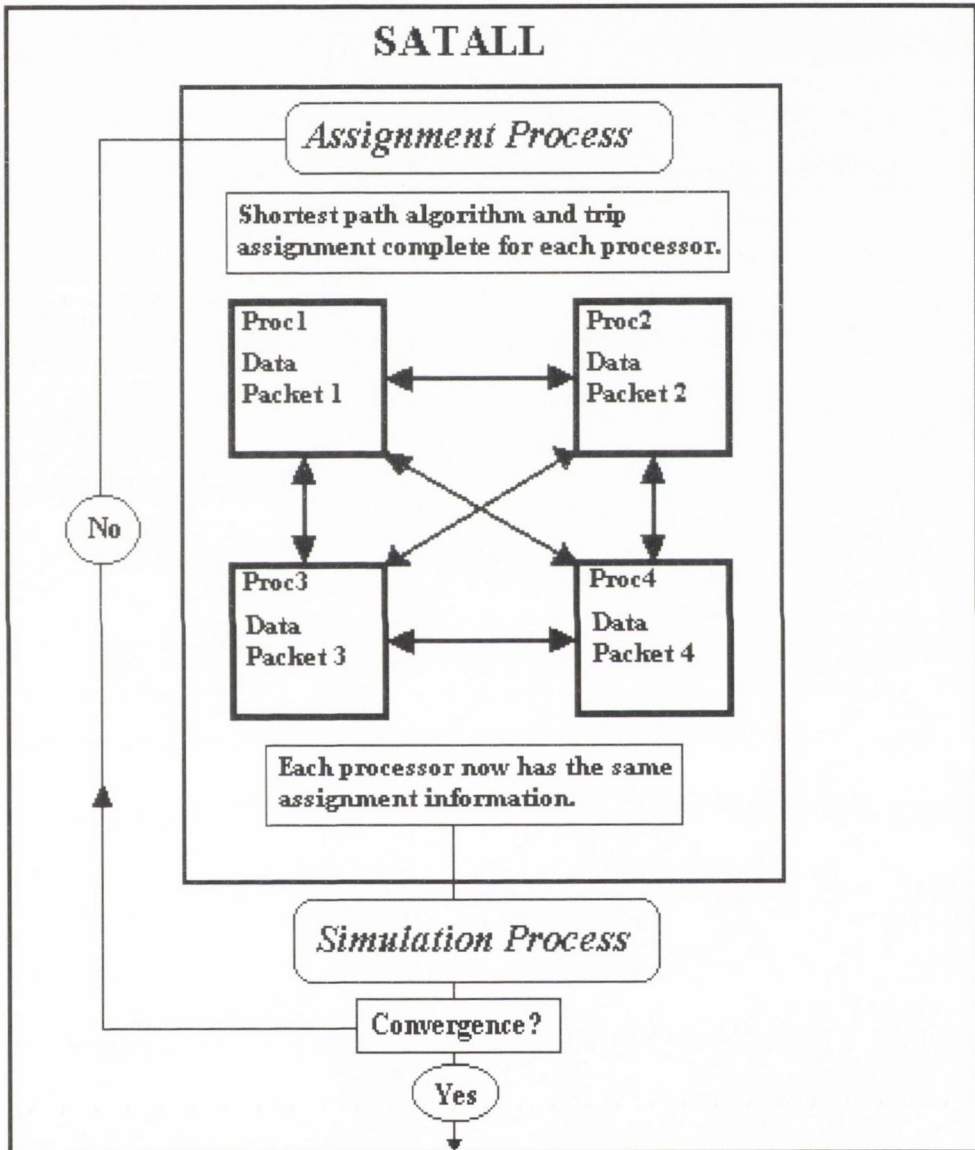
simulation stage. Once the assignment and simulation stages have converged the program outputs the network flows for assessment using one of the SATURN analysis



**Figure 5.2: Simplified diagram of the overall data flow, excluding inter-processor communication, through the computationally intensive program SATALL for each processor involved in a parallel application.**

programs. The Frank-Wolfe algorithm is used many times while the program iterates between the assignment and simulation stages. Also, within the Frank-Wolfe algorithm the shortest path algorithm is used many times per iteration of the assignment stage.

This loop within a loop helps to explain why the shortest path algorithm takes up over 80% of the CPU running time for SATALL.



**Figure 5.3: Simplified diagram of inter-processor communication during the assignment process and the SATALL program running in parallel.**

Therefore, the use of coarse-grained subtasks is profitable in the context of this strategy due to the fact that the areas to be reprogrammed in parallel are substantial in terms of computation and can run in parallel without having to communicate very often with other subtasks also running in parallel. This reduces the inefficiencies associated with

communication overheads, which would be incurred in a fine-grained parallelisation as subtasks running in parallel would have to communicate more frequently; see Chabini (1998) for more detail of fine-grained parallelism in shortest path theory. Creating these subtasks is achieved by dividing the total number of origins by the number of processors being used for the implementation, an early implementation of which is due to Chabini et al. (1997). The result is that there is a very high ratio of ‘computational execution to data input’, which is what allows distribution of the sequential code and a parallelisation approach based around the shortest path loop instead of inside the shortest path algorithm.

Figure 5.3 shows a simplified diagram of inter-processor communication during the assignment process of the SATALL program running in parallel. It is noted that the inter-processor communication only happens at the end of an assignment. After each assignment an inter-processor communication takes place where each processor gathers the network flow information from the assignment on every other processor and, if equilibrium has not been reached or a convergence parameter met, the Frank Wolfe algorithm continues until one or the other has. Once equilibrium has been achieved the SATALL process continues with the simulation process.

This represents a very pragmatic approach, which will give far better performance than fine-grained parallelisation of the shortest path algorithm, which has proved very difficult; see the review in Section 4.7.5 for more detail. In this implementation, each processor in the distributed system will have a copy of the transport network, the trip matrix and a copy of the sequential shortest path algorithm. The parallelisation will result in each processor having a coarse-grained subtask that will be able to find the shortest paths for a subset of origins equal to the total number of origins divided by the total number of processors plus some remainder, and, once each processor has found the shortest paths for its subset of origins, it then assigns the respective flows to this newly found subset of shortest paths based on the O-D demand given in the trip matrix. When this has been completed the shortest path information with the assigned flows, having also been loaded in parallel, are gathered together by each processor from the other processors and reconstituted. The program continues in this fashion for each iteration

of the assignment with the updated network flow information until convergence or a stopping parameter has been reached. In this way any number of processors can be utilised to solve the problem. The data decomposition approach is simply performed for the total number of processors available. This approach avoids the centralised approach described by Hislop et al. (1991), which is limited by the computational bottleneck inherent in a centralised system arising from traffic levels, or the transportation network, becoming too large.

Chabini et al. (1997) implemented two versions of code for the assignment area: one using a distributed memory platform and message passing and the other using a shared memory platform and threads. The drawback of the shared memory version is the limited scalability. Most shared memory systems are more expensive and use fewer processors than distributed systems. At some stage further processors cannot be used as the contention between processors in trying to access the memory at the same time becomes too large. Chabini et al.'s (1997) distributed memory message passing version uses PVM instead of MPI. However, Chabini et al. (1997) and others also opted for a master-slave paradigm whereas a Single-Program, Multiple-Data (SPMD) paradigm is being used here.

The SPMD approach should be more efficient and scale better over multiple processors. The SPMD paradigm is one where each processor runs the same program but acts on a different set of data. The way in which this works is that a duplicate of the parallel program is distributed over multiple processors and each duplicate program runs in parallel, acting on a different set of data and communicating between each other only when altered data is needed by the other processors in order for the parallel application to continue. This communication also occurs at approximately the same time across multiple processors because each processor works on an equal share of the data that has been allocated to it and performs the same tasks on this data, which leads to an efficient load balance amongst processors. Allied to this the MPI interface will be utilised to maximise performance and efficiency from a message passing point of view and allow the parallel solution to be portable across different parallel systems, see Appendix 5 for

a sample of how the MPI code is used in conjunction with new coding to parallelise a subroutine within SATURN.

### 5.3 *Results and Discussion*

#### 5.3.1 *Introduction*

A number of different transport network samples were analysed on three different computing platforms. Comparisons were then drawn between outputs from the parallel model and outputs from the sequential model on each system. The machines that were used were the Intel dual processor i686, the IBM RS/6000 F50 (4-processor) SMP shared memory system and the IBM RS/6000 Scalable POWERParallel SP2 supercomputer, see Appendix 6.

The IBM RS/6000 SP2 System is a distributed memory multi-processor parallel system composed of processor nodes. Each node contains a single Power2SC processor running at 160MHz and runs its own copy of the AIX V4.3.3 operating system. Its peak performance is 640 MFLOPS with 4 FP results/clock, an L1 Instruction Cache of 32 K and an L1 Data Cache of 128 K. The memory configuration is as follows:

- Node 01 >> node 16: 1 GB memory;
- Node 17 >> node 26: 512 MB memory; and,
- Node 27 >> node 48: 256 MB memory.

The total peak performance of the system is 30GFlops. The nodes are connected together using the High Performance Switch Omega Network. For this research 256MB of memory were assigned, as a minimum, to the execution of the program per processor.

The IBM RS/6000 F50 is a shared memory system with 4 PowerPC-604 processors running IBM's UNIX variant, AIX V4.3.3, and using a Symmetric Multi-Processing (SMP) interconnection. Each processor has a peak performance of 332 MFLOPS with an L1 instruction cache of 32 K, L1 data cache of 32 K, L2 cache of 256 KB and an L2 associativity of one.

SATURN, and hence the parallelised version tested here, requires a trip matrix as input and this specifies the number of trips, or journeys, from zone 'i' to 'j' for all zones,  $T_{ij}$ . The trip matrix has been developed by the DTO using different data sources such as vehicle surveys, home interviews, etc. In the case of this research, multiple user classes (MUC), i.e., the private car and heavy goods vehicles (HGVs), are assessed. In the case of MUC assignment involving two user classes, HGVs and private cars, two separate trip matrices are required to describe their movements. To assess these two matrices a 'stacked' orientation is used where the two matrices are 'stacked' on top of each other, making up a single 'unformatted' matrix file (matrix.UFM) representing the trip demand for each O-D pair, before being assigned to the network. This single trip matrix file is input to the model and yields results that refer to the movements of both user classes combined.

The model also requires a transportation network input file (network.UFN) that specifies the physical structure of the road network, and can be thought of as the supply to the demand given in the matrix file. In the case of the DTNM this involves the definition of both a simulation network and a buffer network. The simulation network is coded in more detail than the buffer network and defines node and link based data. The node-based data includes traffic signals, priority junctions and roundabouts, while the link-based data describes the roads between intersections (nodes). The type of detail modelled at this stage includes restricted movements, banned turns, one-way streets, link and junction capacities, travel times, link distances, cost-flow curves, priorities at junctions, gap acceptance parameters and saturation points, amongst others. The buffer network lies outside the simulation network and is programmed in less detail but accounts for a larger area of less importance. The essential difference between the two networks is that the simulation network models both link and junction data while the buffer network models only link data.

Both the trip matrix and road network are used as inputs to the route choice model, which allocates (assigns) trips to routes through the road network. The GDA trip matrix has 432 fine zones, with finer zones in the City Centre and coarser zones further out, giving 432 sources (O-Ds). The highway network is made up of a simulation network

and a buffer network, combining to give over 1,220 simulation nodes (junctions) and 3,900 links (connections between junctions) throughout the GDA.

The model uses up approximately 1.5 MB of space when running on each processor; this was found using the AIX command 'TOP'. Section 5.3.2 assesses the convergence of the parallel model using the full GDA road network with the associated trip matrix. Section 5.3.3 discusses results from the assessment of the accuracy of the parallel model. Finally, Section 5.3.4 details the performance results, discussing the parallel model in relation to the speed-up, efficiency and time-savings when compared to the sequential implementation.

### 5.3.2 Convergence

The parallel model was run on the 'IBM RS/6000 SP2' and 'IBM RS/6000 F50' HPC parallel platforms and sequentially over a single processor. Convergence was reached after the same number of iterations for all model runs. Once it was clear that the parallel model converged the output was compared to that of the sequential SATURN model. ISTOP is a percentile convergence parameter that stops the loop between the simulation and assignment from continuing if 'ISTOP' percentage of the link flows change by less than a predefined value, 5% in this case. After the 15<sup>th</sup> loop the following output is produced for the parameter ISTOP where ISTOP is set equal to 90%.

- For the sequential SATURN model run on the IBM ibix machine, using one processor, 89.8% of the assigned flows are within 5% of their values from the previous simulation/assignment loop iteration.
- For the i686 dual-processor machine, using two processors, 90.7% of the assigned flows are within 5% of their values from the previous simulation/assignment loop iteration.



- For the IBM ibix machine, using four processors, 91.3% of the assigned flows are within 5% of their values from the previous simulation/assignment loop iteration.

This demonstrates that the parallel model converges with at least the same measure of accuracy that the sequential SATURN model is able to achieve. This has been found to hold for any number of processors that have been used, up to and including 16 processors. It is noted that ISTOP finishes at a slightly different percentage for each of these tests. In the first case above, using the IBM ibix machine, the model gives results where 89.8% of the assigned flows are within 5%. This compares to 90.7% and 91.3% for 2 and 4 processors, respectively.

The slight differences in the convergence results are not considered large enough to represent inaccuracy in the conversion of the parallel model for a number of reasons:

- The first and primary reason is that the parallelisation of the sequential model does not change the algorithm itself but rather inserts a parallel loop around it. Therefore, as the convergence theory remains unchanged for the parallel model application the convergence within the parallel model is unaffected;
- The slight changes in the results shown above are not unusual, i.e. slight changes can result when running the sequential model using the same input data, which means that it is the inherent nature of the model itself and the algorithms used within it, and not the parallelisation process, that produce slightly differing results from test run to test run; and,
- The parallelisation process may result in slight differences in rounding errors associated with the processes of the model but these would be so small as not to be of any significance.

### 5.3.3 Comparison of Model Outputs

Initially a small transport network sample was used to determine whether the parallel model was running accurately, i.e., that the parallel model was not producing spurious outputs. This sample has the parameter MASL (**M**ultiple **A**ssignment **S**imulation **L**oops) set equal to three, where it would usually be set equal to a high number such as 30 in the case of the DTNM. MASL controls the number of assignment/simulation loops so when MASL is set equal to three the program is stopped after three loops. The resulting output from the sequential model over a single processor and the parallel model run using four processors on the IBM ibix machine differ by less than 1%, as demonstrated in Table 5.1.

The difference in results is due to the same reasons given in Section 5.3.2 above, which are the inherent nature of the sequential model itself and the algorithms within this model and due, to a lesser extent to the rounding error associated with the complex nature of the calculations involved in the assignment/simulation processes and iterative loop, and, the convergence of this loop to equilibrium. In practice the accuracy of the model output has been found to hold for any number of processors that have been used, up to and including 16 processors on the IBM SP2 HPC.

**Table 5.1: Sequential SATURN output compared with parallel model output.**

	SATURN (sequential, single processor)	IBM ibix (shared memory, four processors)
<b>TRANSIENT QUEUES</b> (PCU. HRS./HR.)	5970	5987
<b>OVER CAPACITY QUEUES</b> (PCU. HRS./HR.)	24821	24638
<b>LINK CRUISE TIME</b> (PCU. HRS./HR.)	16231	16231
<b>TOTAL TRAVEL TIME</b> (PCU. HRS./HR.)	47022	46856
<b>TRAVEL DISTANCE</b> (PCU. KMS./HR.)	837029	837050
<b>OVERALL AVERAGE SPEED</b> (KPH)	17.8	17.9

Once the accuracy and convergence of the parallel model were determined to be acceptable the models performance was assessed for the DTNM.

### 5.3.4 Performance

The performance of the HPC system, combining a software element in the parallel algorithm and a hardware element in the HPC platform, is assessed. A detailed review of performance analysis and the performance parameters used for this research is provided in Section 4.7.4. However, the definition of performance depends on the reason behind the use of HPC theory in the first instance. In general a reduction in the computation time of the program, or model, is sought. Hence, the improvement in performance is measured by comparing the sequential time to run the program on a single processor against that of the parallel application across a number of processors.

The parameters used include the measure of speed-up,  $S(n)$ , defined as the ratio of the elapsed time, or serial time, when executing a program on a single processor,  $T_s$ , to the execution time when a number of processors,  $n$ , are available,  $T(n)$ . Therefore;  $S(n) = T_s / T(n)$ . The efficiency,  $E(n)$ , is defined by Eager et al. (1989) as the average utilisation of the 'n' allocated processors. Therefore, the relationship between efficiency and speed-up, described by Eager et al. (1989) as the average processor utilisation, can then be defined by  $E(n) = S(n) / n$ , leading us to the definition of linear speed-up, i.e., where the efficiency remains at 1 while the number of processors increases beyond two. In practice an efficiency of 1 is impossible to maintain when the number of processors increase, even for software whose code is designed specifically for HPC parallel platforms.

Gurd et al. (1985) found that programs with a similar value of average parallelism exhibit virtually identical speed-up curves and that the higher the value, the closer the program got to achieving 100 % utilisation, i.e. an efficiency of 1, on a graph. Eager et al. (1989) define the average parallelism in a number of ways including the following two:

- The average number of processors that remain busy during the execution of a software system given an unlimited number of available processors; and,
- The speed-up given an unlimited number of available processors, i.e., the maximum possible speed-up.

'Benefit verses cost' profiles, describing the relationship between the benefit of a scheme and the costs associated with that scheme, are then discussed with particular reference to the efficiency – execution time profile and the 'knee' in such a profile, see Denning (1980) and the discussion in Section 4.7.4. It is shown that the location of the 'knee', i.e., the number of processors used at this location, is well approximated by the average parallelism. Further to this, the 'knee' occurs where the ratio of efficiency to execution time,  $E(n) / T(n)$ , is maximised, where 'n' is the number of processors allocated to the computation at this point. Eager et al. (1989) prove that at the 'knee' of the efficiency – execution time profile a speed-up,  $S(n)$ , and efficiency,  $E(n)$ , of at least 50% of their maximum is guaranteed. Therefore, the 'knee', or rather the number of processors used to achieve the location of the 'knee' on the efficiency – execution time profile, appears to give the optimum number of processors to apply to a parallel model in order to achieve the best balance between efficiency and execution time.

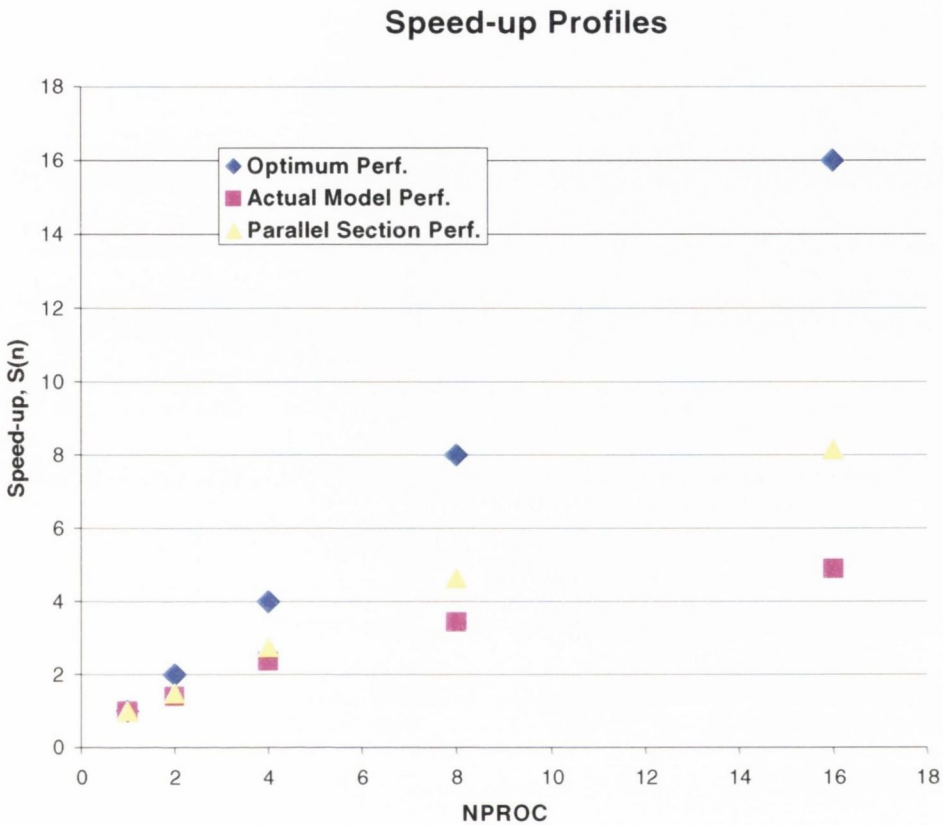
These performance parameters are reported in Table 5.2, which gives the execution time taken to complete the parallel model on the SP2 machine, with and without the sequential section, the speed-up,  $S(n)$ , the efficiency,  $E(n)$ , and, the ratio of efficiency to execution time,  $R$ , of the parallel model and of the parallel section by itself. Results shown for the total model,  $T(n)$ , where 'n' = 1 also use the parallel algorithm.

A graph of speed-up verses the number of processors being used (NPROC) for the parallel implementation and for the parallel section of the new parallelised model can be seen in Figure 5.4. The graph also shows the theoretical maximum, where there is a linear relationship between the number of processors used and the resulting performance gain, shown as the "Optimum Performance". This graph was produced by solving the GDA trip matrix assignment/simulation problem using 1, 2, 4, 8 and 16 processors. More processors were not used because of administrative limitations due to large demands and a heavy workload on the SP2 platform.

**Table 5.2: Performance parameters including the execution time, speed-up, efficiency and the ratio of efficiency to execution time.**

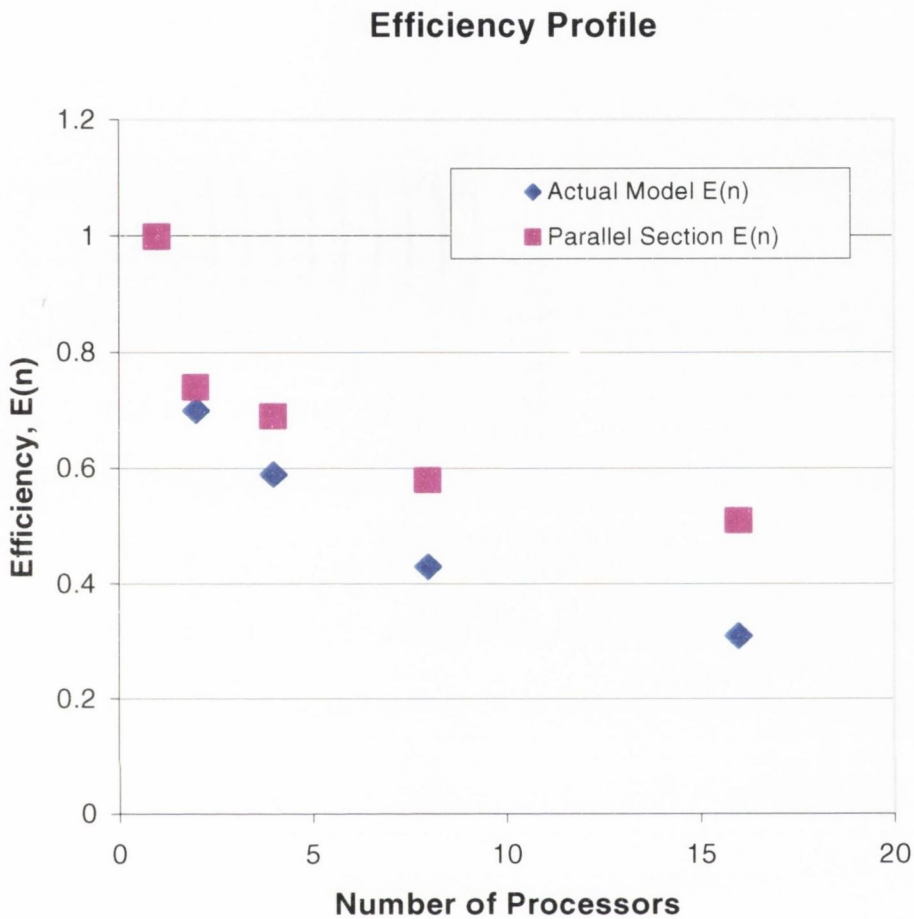
	<b>N, the number of processors used to run the model</b>				
	<b>1</b>	<b>2</b>	<b>4</b>	<b>8</b>	<b>16</b>
<b>T(n) of total model</b>	115.25min	82.1mins	48.73mins	33.66mins	23.63mins
<b>T(n) of parallel section</b>	105.05min	71.14mins	38.13mins	22.53mins	12.87mins
<b>Sequential segment</b>	10.2mins	10.96mins	10.61mins	11.14mins	10.77mins
<b>S(n) of total model (seq. incl.)</b>	1	1.4	2.36	3.42	4.88
<b>S(n) of ass. and s.p. (parallel section)</b>	1	1.48	2.76	4.66	8.16
<b>E(n) of total model (seq. incl.)</b>	1	0.7	0.59	0.43	0.31
<b>E(n) of ass. and s.p. (parallel section)</b>	1	0.74	0.69	0.58	0.51
<b>R of total model</b>	0.0087	0.0085	0.012	0.0127	0.0129
<b>R of parallel section</b>	0.0095	0.01	0.0182	0.0259	0.0396

It is noted from Table 5.2 and Figure 5.4 that while the speed-up for the parallelised model drops off quickly with a value of 4.88 for 16 processors, the performance of the parallel section of the new model is far better with a value of 8.16 for 16 processors. The obvious reason for this is that the sequential section of the parallel model takes the same amount of time to complete no matter how many processors are being used. Therefore, the relative importance of that sequential section increases as the number of processors being used increases, which was predicted beforehand. The performance



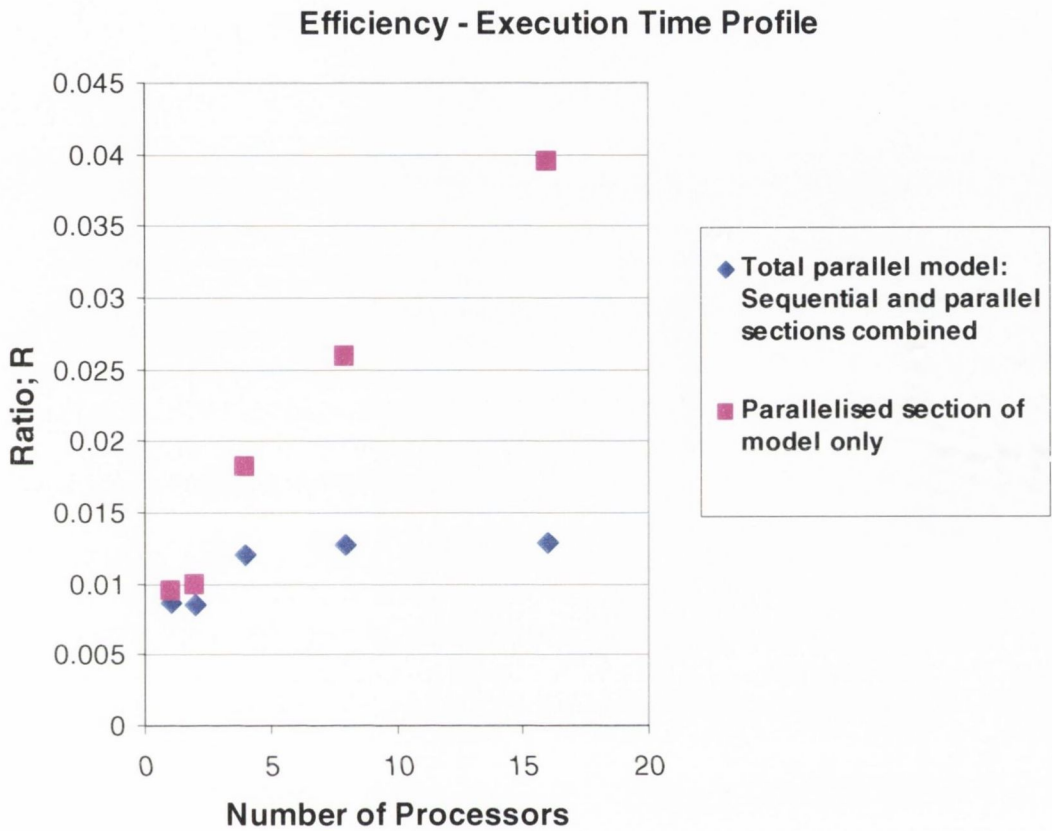
**Figure 5.4:** Speed-up versus the number of processors used for the optimum performance, for the actual performance of the model, and, for the parallel section performance compromising the shortest path problem and assignment of flows.

speed-up of 8.16 is more than that achieved by Greenwood and Taylor (1993), which was approximately 7, see Section 4.7.4.2. This is reinforced by examining the efficiency of the parallel model, see Figure 5.5. The efficiency,  $E(n)$ , of the parallel model decreases quickly as the relative importance of the sequential section increases with the number of processors used. The sequential section represents 9% of the execution time,  $T(n)$ , of the model using 1 processor but it represents 54% when 16 processors are used. As such the sequential nature of the model represents the dominant limiting factor in this parallel implementation.



**Figure 5.5:** Graph showing the relationship between efficiency and the number of processors used for the total model with the sequential section included and for the parallelised section of the model only.

It is also noted that the efficiency,  $E(n)$ , execution time,  $T(n)$ , profile of the parallel model, represented by the ratio of one to the other,  $R$ , while still increasing for 16 processors at 0.0129, appears to be levelling off quite quickly, see Table 5.2. This is demonstrated clearly in Figure 5.6, where the ratio,  $R$ , levels off quickly for the parallelised model. This is expected due to the sequential section inherent in this parallel application, however it does suggest that the ‘knee’ of the model, as described earlier, lies somewhere around 0.013, the exact position of which would be found for the model tested using an increased number of processors. This point gives a reasonably accurate location for the average parallelism of the parallel model as described earlier.



**Figure 5.6: The ratio,  $R$ , measuring the efficiency,  $E(n)$ , execution time,  $T(n)$ , profile for the total model with the sequential section included and for the parallelised section of the model only.**

The performance of parallel applications is usually measured by assessing the parallelised section itself, i.e., excluding any sequential element associated with inherently sequential algorithmic content or the sequential nature of I/O. In this respect, the efficiency,  $E(n)$ , of the parallel section, comprising the shortest paths problem and the loading of flows to these shortest paths, decreases far slower than the parallel model as expected, see Table 5.2 and Figure 5.5 for a comparison between the parallel model efficiency and the parallel section efficiency. Using 16 processors the efficiency of the parallel section is 0.51 compared to 0.31 for the parallel model. Additionally, the ratio,  $R$ , is increasing with additional processors leading to the conclusion that the number of processors needed to reach the ‘knee’ for the parallel section is far more than is available for this research.



Examining Figure 5.6 it is noted that for the parallel section, after an initially poor increase in the ratio,  $R$ , from 1 to 2 processors, the efficiency – execution time profile recovers and an almost linear relationship is noted for 4, 8 and 16 processors. The rate of decrease of the efficiency is reducing and the ratio,  $R$ , suggests that the pattern of the graph will continue for a considerable number of extra processors. As such, it is hard to establish the effect of communication overheads on the parallel section of the model without actually applying the model in practice to more processors. Future work could, therefore, concentrate on finding the average parallelism of the parallel section of the model and reducing the time spent in other sequential areas of the model.

Using a direct comparison of the times taken for completion of the parallel model with different values of NPROC, the number of processors utilised, it can be seen that the SP2 HPC platform is also far more efficient even using just one processor, see Table 5.3. The times shown in Table 5.3, in minutes, give the total time taken for the model to run to completion using three different machines over 1, 2, 4, 8 and 16 processors. While the processor clock speeds may vary slightly for these different parallel systems the performance of the SP2 platform is mainly due to the amount of Random Access Memory (RAM) available on the SP2 machine, which far exceeds that of the shared memory Ibis machine or the dual-processor Intel machine. The results demonstrate the portability of the parallel model across different parallel systems.

**Table 5.3: The time taken to complete the parallel model, with the sequential section and I/O included, for the three systems analysed.**

	<b>NPROC, the number of processors used to run the model</b>				
	<b>1</b>	<b>2</b>	<b>4</b>	<b>8</b>	<b>16</b>
<b>SP2 (multi proc.s)</b>	115:45mins	82:10mins	48:53mins	33:38mins	23:40mins
<b>IBM Ibis (4 proc)</b>	171:55mins	101:1mins	63:25mins		
<b>Intel (2 proc)</b>	163:05mins	95:56mins			

Using just one processor the SP2 machine takes approximately 115 minutes as compared to that of 172 minutes and 163 minutes for the Irix and Intel computers, respectively. Using 8 or 16 processors the time is 3.4 and 4.9 times faster than a single processor, respectively, and the time of 23.66 minutes for the SP2 platform to complete the model run using 16 processors saves 92 minutes, or 80%, when compared to the time taken for a single processor to execute. This also represents large savings when compared to the time taken for a single processor to finish the job on the Irix or Intel platforms.

## *Chapter 6*

# *Inner City Congestion Charging in the Greater Dublin Area*

## **6.1 Introduction**

A package of measures has been outlined by the Dublin Transportation Office (DTO) to offset the current imbalance between transport demand and supply on the typically heavily congested roads of the Greater Dublin Area (GDA) highway network. As part of this package, traffic demand management was proposed as an important method for the control of traffic congestion.

In Chapter 2 it was recognised that due to the institutional and regulatory inefficiencies, the cumbersome and time consuming planning system, and, the high economic growth since the DTI began, delays to infrastructural projects became inevitable and unrealistic planning targets were set. Consequently, a substantial transportation deficit has arisen in the GDA. In response, it has been recognised by the DTO that a comprehensive study is necessary to develop a demand management strategy to reduce growth in overall travel by private car and to effect further modal transfer from private car to public transport modes.

Demand management seeks to reduce the growth in travel while maintaining economic progress and is designed to encourage a transfer of trips, especially during peak periods, from the private car to sustainable modes of transport (DTO, 2000). Road user charging is an important tool in demand management strategy. There are a number of established forms of road user charging including cordon charging, area licensing, screen-line tolling and pricing by time or distance; see the literature review detailed in Chapter 3.

Chapter 3 describes the development of road user charging and briefly introduces important elements of design and implementation. The political and social aspects of acceptability are then examined, followed by a review of the effectiveness of road user charging. A number of studies of pre and post implementation cases are then examined and price elasticities of user demand are considered in this context. Finally, conclusions are drawn from this review of international literature and practical experience, making particular reference to the research carried out for this thesis. A summary of the relevant conclusions from this Chapter is presented in Section 6.2.

This Chapter reports on the impact of a series of road user charging inner city cordon scenarios. These scenarios are tested over the GDA highway network, used in the DTNM, and using the base year trip demand input matrix, also produced for the DTNM by the DTO. Preliminary modelling was carried out on the original 1997 base year trip matrix and network files obtained from the DTO, see Appendix 7 for a summary of the results from these tests. Modelling was then carried out for the 2000 and 2006 trip matrix and network definitions with particular reference to the 2006 model; Appendix 8 and Appendix 9 summarise the results from this testing, respectively. The parallel model was used to carry out these modelling tests; the details of this parallel implementation are reported in Chapter 5. The parallel model executes on the IBM SP2 supercomputer described in Section 5.2.3, which allows for parallel processing of model tests while retaining accurate modelling behaviour.

Section 6.2 briefly summarises the transportation background, which has been dealt with in some detail in Chapter 2. The cordon scheme location is also discussed with conclusions drawn on the best location for an inner city congestion charging cordon. Section 6.3 presents the methodology. This describes the natural geographical location of the cordon and its size and position relative to the GDA. A brief review of previous demand management studies in the GDA is then discussed in Section 6.3.2. Section 6.3.3 examines the input data files for the cordoned model implementations and discusses the file content and structure. Section 6.3.4 summarises the conclusions reached from the literature review of road user charging in Chapter 3, with particular reference to the charge levels and the range of price elasticities of user demand in response to these charges. The elasticity demand function, which relates elasticity to charge, trip demand and generalised cost, is then explained and generalised cost is summarised in Sections 6.3.5 and 6.3.6, respectively.

Section 6.3.7 presents more general information on the inner city cordon scenarios modelled. Differences between the highway networks of 2006 and 2000/1997 are explained and a discussion on elastic assignment and the attempts of this research to include elastic effects over a fixed demand transportation network model are explained.

The test scenarios are detailed in Section 6.3.8, followed by the results and discussion in Section 6.4.

The modelled cordon scenarios are described and discussed using a range of transportation variables including travelled distance, travelled time, average speed and queuing. Some details are also provided on the effects of the cordon scenarios on fuel consumption over the entire highway network as a whole.

## **6.2 Background**

### *6.2.1 Transportation deficit*

The DTI began a transport planning process in 1988 permitting regular review and assessment of the impacts of land use and other policies on transportation and vice versa. In 1992, Phase 2 of the DTI began resulting in the production of the DTI Strategy in 1994 which laid out transportation plans for Dublin. The DTI produced a number of possible land use development scenarios with different growth levels for the future, see Section 2.4. The objectives of the DTI Strategy were to encourage economic regeneration and development and to help maintain and reinforce Dublin City centre as the country's prime commercial and cultural centre, See Section 2.4 and Appendix 1 for more details.

The DTI Strategy could never provide a definite long-term solution because it was based on a number of future land use scenarios. A land use strategy would have to be designed to complement it, which was recognised in 1999 with the production of the SPG. The 4-stage transportation modelling approach was used to produce the DTNM, which uses trip generation and attraction, trip distribution, modal split and trip assignment (using the SATURN traffic simulation and assignment model), see Section 2.4.4.

Forecasts for population and traffic growth made for the DTI Strategy have been very conservative, actual trends have outstripped all expectations, i.e., predicted population for 2001 was exceeded in 1997 and the level of demand for travel has increased by

double what was predicted for 2001 in the 1991 DTI forecast. The institutional and regulatory inefficiencies; the cumbersome and time consuming planning system; and, the high economic growth since the DTI began, have also resulted in delays to infrastructure projects and unrealistic planning targets, see Section 2.5.

Any increase in person-trips by car over that predicted by the DTI is defined as a transportation deficit (DTO, 1998). Two main problems exist as the main causes of the substantial transportation deficit in the GDA and these can be summarised as follows:

- The first is that there has been considerable 'slippage' in the implementation of the DTI Strategy, although with the emergence of the DTO to oversee its implementation, progress has been made; and,
- The second is that there has been enormous growth in the economy resulting in very large increases in the level of traffic and increased congestion. This economic growth was not predicted in the DTI Strategy. Demographic and economic indicators for the GDA are presented in Table 6.1.

The high rate of economic growth has resulted in a large increase in the demand for travel. Peak hour trips between 1991 and 1997 have grown by 45% or 78,000, 71,000 of which are due to private car commuting. The average journey time has increased from 31 minutes to 43 minutes over the same period with 72% of peak hour trips being accounted for by the private car.

The DTNM was produced as part of the DTI to make transportation forecasts. The DTO have developed this model further to produce new forecasts for the GDA (DTO, 2000). A percentage growth factor per year to predict future year trip matrices is not used. Instead, land-use planning and employment information on a zone-to-zone basis is used to predict the growth for each zone within the GDA. The forecast produced by the DTO for 2016 predicts an even worse outlook with a 95% increase in the total peak hour trips over and above the predicted demand in 1997, while the off peak demand for trips in 2016 will also exceed the peak period demand in 1997. If, however, other

strategies of transportation management were to be used, such as demand management e.g. road user charging, this situation could potentially be improved.

**Table 6.1. Factors Influencing Traffic Growth (DTO, 2000)**

<b>Greater Dublin Area</b>	<b>1991</b>	<b>1996</b>	<b>1999</b>	<b>2016</b>
Population (million)	1.35	1.41	1.46	1.75
Households ('000)	402	446	521	675
Employment ('000)	452	549	602	878
Unemployment rate	16%	12%	6%	5%
Car ownership (per 1000 pop.)	247	292	342	480
% growth in GDP since 1991	-	42%	79%	260%

In 1999, the SPG, a “*framework for integrated land use and transportation for the sustainable development of the GDA up to the year 2011*”, was produced, see Section 2.6, which provided the first co-ordinated settlement plan for the GDA and the starting point for the latest transportation plans. From a transportation perspective, the SPG objectives are to consolidate development and increase density in the GDA so as to provide a basis for a better public transport system, and, to facilitate a shift away from the private car to public transport. The SPG recommend that demand management measures should be incorporated into all strategic planning and implementation to reduce traffic demand, dissuade private car commuting and encourage the use of public transport, see Section 2.6.1.

The DTO 2000-2016 Strategy also advocates an integrated transportation strategy involving demand management and infrastructural and service improvements, see Section 2.6.2. It is recognised that a comprehensive study is necessary to develop a demand management strategy to reduce growth in overall travel by private car and to effect further modal transfer from private car to public transport modes. Likely benefits of implementation of the DTO Strategy include between 27% and 57% less congestion than if the Strategy was not implemented. Other benefits include reductions of 41%,



34% and 35% in energy consumption, emissions and accidents, respectively (DTO, 2001).

### 6.2.2 *Road user charging*

It has been demonstrated in Chapter 3, Section 3.5, that road user charging can substantially affect user behaviour and hence, traffic behaviour. Based on modelling studies and practical implementations it has been shown that road user charging can be used to raise revenue, reduce environmental pollution and reduce congestion. With the reduction of car trips during peak periods a transfer of trips to other modes can be expected as the primary response to a road user charging scheme.

Locational impacts appear to be very small, leading to the conclusion that geographical layout plays only a minor role in the overall effectiveness and impact of road user charging schemes. However, from a business perspective congestion charging of a city centre tends to increase the number of shops within the city centre, as opposed to a decrease in the number as was initially feared.

Though international experience of road user charging is, in general, positive, pitfalls are especially prevalent when dealing with political and social acceptability of congestion charging schemes, i.e., the experiences in Stockholm, Hong Kong and Cambridge. With respect to the implementation of congestion charging schemes a city centre cordon scheme with a fixed charge appears to be favourable, especially in relation to social and political acceptance. Once a simple scheme such as a city centre cordon has been introduced then it is far more likely that the general public, and hence politicians, will accept a more complex upgrade like that implemented in Singapore.

It was found in Chapter 3, Section 3.4, that in order to implement a successful charging scheme it is, therefore, necessary to design the scheme not just to be effective in achieving its objective but to be acceptable to the general public as well. Further to this, Jaensirisak (2002) has found that acceptable road user charging schemes can be designed to be both acceptable and effective by limiting the area of charge to within the city centre and having a fixed charge per day. It was also noted that support would

increase significantly if there was substantial improvement to the environment. Therefore, any road user charging scheme must be implemented as part of an integrated strategy that combines transport demand management with infrastructure and service improvement as suggested in the SPG (1999) and by the DTO (2000, 2001).

From the study of road user charging studies, pre and post implementation, it is concluded that the design and implementation of congestion charging schemes is very important for public acceptability, particularly with respect to charge levels, enforcement, privacy and the degree of public understanding. Political acceptability is difficult to attain because most users tend to lose out. One approach for overcoming people's suspicion of change is to apply a strategy of incremental development (Small and Gomez-Ibanez, 1998). For example, the Norwegian toll rings began as a means of raising revenue for financing regional transportation infrastructure but have begun to incorporate traffic management as a subsidiary goal.

At a more advanced level it is evident that introduction of a simple congestion charging cordon scheme in the first instance is preferential if the scheme is to be a success (GOL, 2000; Oscar Faber et al., 1998). More complex systems are more likely to be successful after a simple system has been seen to work by the general public. In Singapore a very simple congestion charging cordon scheme was introduced first and has since been developed into a more complex system with multiple cordons and screen-lines over a wider area.

### *6.2.3 Inner city congestion charging cordon scheme*

This thesis reports on research of an inner city congestion charging cordon scheme for the GDA. The inner city location and the cordon implementation of road user charging with a fixed charge were chosen for a number of reasons:

- The recommendations made by Oscar Faber et al. (1998) that the most suitable place to implement an inner city congestion charging scheme would be along the inner city canals and North Circular Road;

- The conclusions reached in the literature review of road user charging, see Chapter 3 and the summary in Section 6.2.2 above, that the best type of road user charging scheme to implement from a social and political standpoint, and, from an implementation and effectiveness standpoint, is a simple inner city cordon scheme with a fixed price;
- Research has shown, see Chapter 3, that the overall impacts associated with the geographical location of the cordon are small in comparison with its effectiveness in reducing congestion and actually attract more business to inner city leisure and retail businesses, which make up the bulk of inner city commercial enterprise;
- The fact that the geography of the city lends itself to a natural cordon around the inner city along the lines of the Grand Canal, the Royal Canal and the North Circular Road, which encircle the inner city in a natural cordon, see Figure 6.2; and,
- Previous implementations of successful congestion charging schemes on an area-wide basis have all used an inner city congestion charging cordon scheme to begin with, e.g., Singapore and London.

The relative location of the inner city cordon to the rest of the GDA can be seen in Figure 6.1. The cordon itself takes up a small portion of the simulation area of the DTNM, which is in turn combined with the larger buffer area lying outside this to produce the overall DTNM highway network that this research models. A larger scale map of the location shows the natural geographical location along the routes suggested above, see Figure 6.2. Figure 6.2 also shows the outer ring to the left of the inner city, which represents the M50 Motorway and the boundary of the simulation area shown in Figure 6.1.

### 6.3 Methodology

#### 6.3.1 Introduction

The inner city road user charging cordon scheme is applied to the DTNM, which uses a

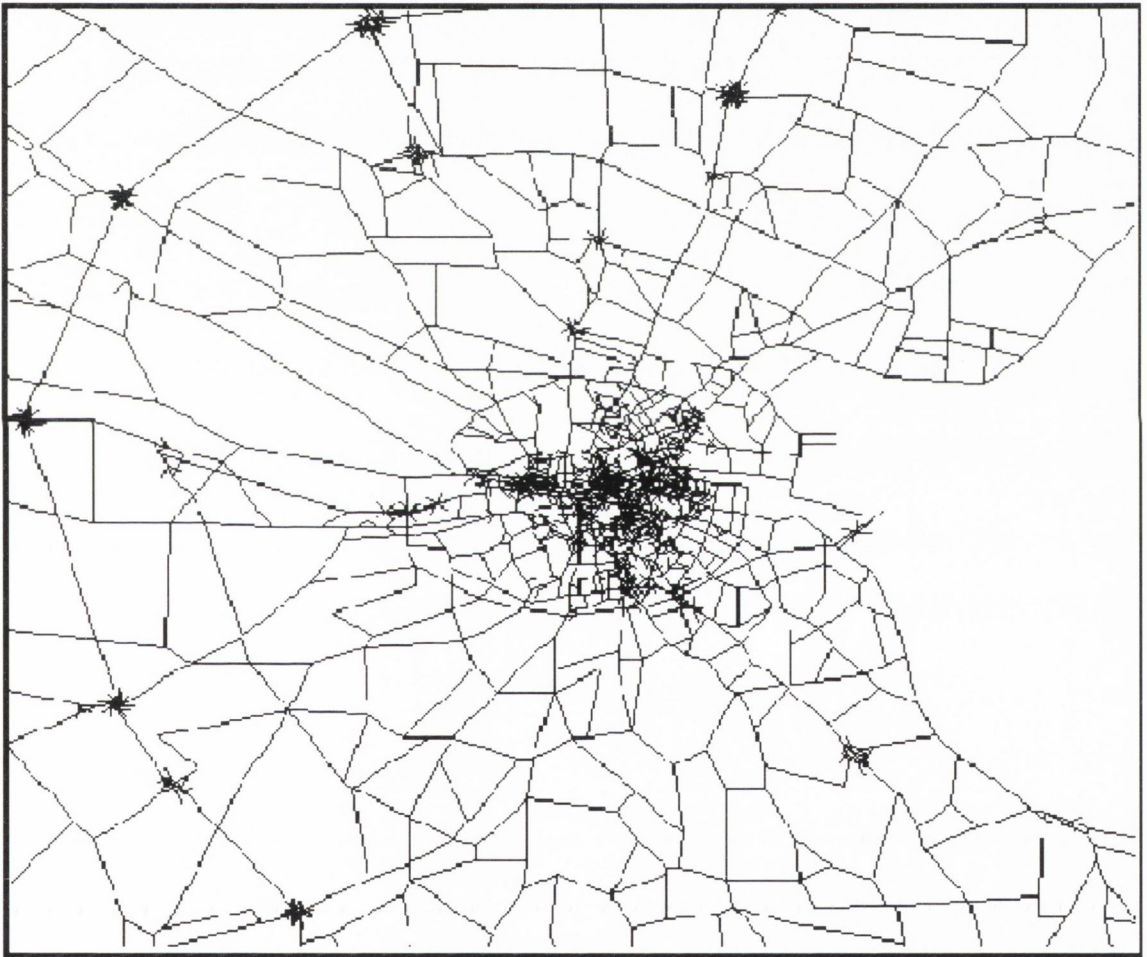


**Figure 6.1.** Basic map of the Greater Dublin Area showing the Metropolitan Area, the simulation area boundary used in the DTNM and the inner city cordon scheme location.

base year trip demand matrix and a highway network (supply) to define the geographical and structural layout of the GDA highway network and the trip demand across that highway network. The modelling will assess the effect of the inner city cordon scheme, whose geographical location lies along the lines of the Royal Canal, Grand Canal and the North Circular Road, see Section 6.2.3, on the modelling

representation of private car and HGV traffic movement provided by the DTO. The cordon scheme will be applied during the morning peak period, from 8am to 9am, using the 1997, 2000 and 2006 highway network and trip demand definitions, also produced by the DTO.

In order to do this, the base year trip matrix and highway network definitions for 1997, 2000 and 2006, must be altered to take into account the inclusion of the inner city



**Figure 6.2: Map showing the natural geographical cordon location around the inner city area and the M50 Motorway.**

cordon. These alterations are made to the input files for the SATURN model used to assign the trip demand to the highway network and simulate the traffic movement and distribution across that highway network. However, in this case the parallel version of SATURN is used instead of the sequential one used by the DTO. The process of

altering the trip demand matrix and the highway network input files is detailed in Section 6.3.3.

Section 6.3.2 briefly summarises the conclusions of the only other road user charging study of the GDA produced by Oscar Faber et al. (1998). There are a number of factors and constraints that must be considered for the road user charging scenarios assessed. Section 6.3.4 details the charge levels to be used in the modelling process and the range of price elasticity's of user demand in response to these charges. Section 6.3.5 looks at the elasticity demand function used in the modelling process to describe the relationship between the number of trips travelling along a route, the charges associated with travelling along that route, the cordon charge to be applied along the route and the elasticity response to that extra charge. Section 6.3.6 details the generalised cost equations used to calculate the charge associated with travelling along a route on the highway network. These costs are used to produce a cost matrix representing each origin-destination (O-D) route in the same way that the trip demand matrix gives the number of trips travelling along each O-D route. Section 6.3.7 describes the cordon tests while Section 6.3.8 details the test scenarios.

### *6.3.2 Previous studies of road user charging in the GDA*

'A Study of Road Pricing in Dublin' was produced by Oscar Faber et al in 1998. This is the only report that has been produced to assess road user charging on an area-wide basis in the GDA, and, in essence was a feasibility study. It examined a small number of demand management scenarios to try to assess the feasibility of road user charging and other fiscal instruments, such as parking charges, for the management of transportation and traffic demand in the GDA. The study also looked at the different forms of road user charging and their implementation and went on to confirm the potential of road user charging in contributing to the management of traffic in the GDA.

The study was undertaken strictly as a preliminary study to assess the potential of road user charging and it was assumed that further information and detail would be required to fully evaluate charging schemes in the GDA. Based on the positive preliminary findings of the study the report recommends that further work should be undertaken to

investigate the feasibility of introducing a road user charging scheme, and in particular a cordon charging scheme, to the GDA highway network.

### 6.3.3 *Cordon modelling input data*

There are two sets of data to be used as input to the model:

- (1) The first data set describes the highway network, including the infrastructure, traffic signals, geographic layout, etc., with over 2500 links, 4500 simulated turns and 320 traffic signals. The data describing the highway network has been provided by the DTO and was initially produced for the DTI, see the DTI Final Report (1994) for more detail; and,
- (2) The second data set is the trip demand matrix, which in the case of the DTNM, is made up of 432 O-D pairs making a 432\*432 matrix. Each element of the matrix is equivalent to the number of trips either originating in or destined for a certain zone, of which there are 432 that make up the study area, which was also provided by the DTO. Initially a 367\*367 matrix was produced using a number of surveying and interviewing techniques and census data as part of the DTI Strategy (DTI, 1994) and was then developed further by the DTO to a 432\*432 size matrix.

#### **6.3.3.1 Model input data: the highway network**

The GDA transport network is made up of two networks, which are better known as the simulation network, which includes the Central Business District (CBD), and the buffer network, the area outside the simulation network. The simulation network encapsulates the inner city, CBD and surrounding urban environment inside the M50 Motorway, and is programmed in more detail than that of the buffer network, see Figure 6.1 for the simulation boundary location. The buffer network represents the greater part of the GDA (other than the simulation area) and also incorporates traffic entering the area along major routes from the rest of the country. The simulation and buffer networks

combined give the complete highway network, as used in the Dublin Transportation Network Model (DTNM).

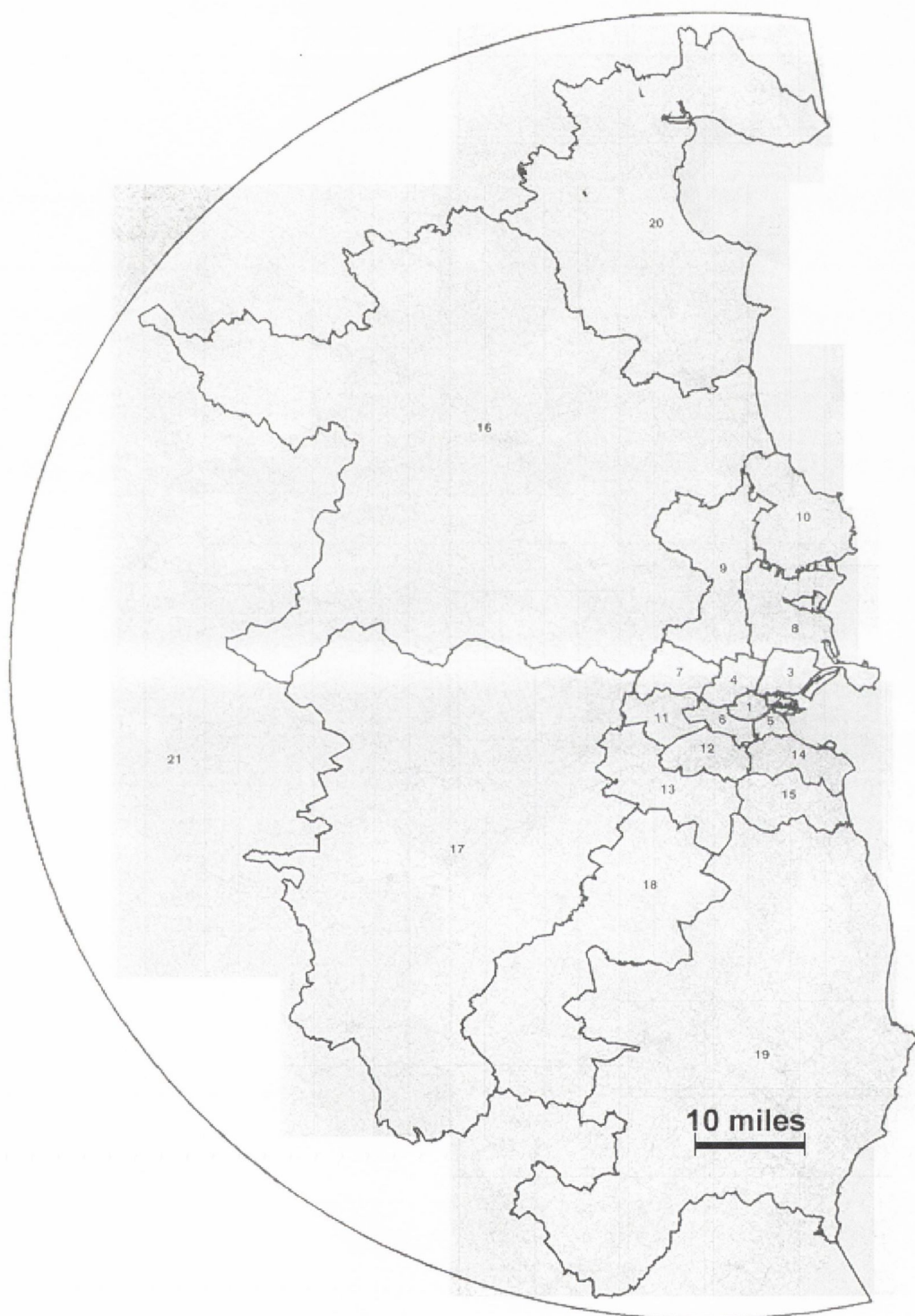
The highway network information is input to the transport model in the form of an 'UnFormatted' Network (UFN) file, which will hence forth be known as the Network.UFN file. The Network.UFN file is produced using another program from the SATURN Suite called SATNET. SATNET takes the information produced in a standard DAT file and converts this text-formatted file to the Network.UFN file format, which is then input to the assignment and simulation model. An example of one of the DAT files used to create a Network.UFN file can be found in Appendix 10.

The network data file is made up of a number of smaller sections including:

- The network title;
- Parameter specifications
- The simulation network and simulation centroid connectors;
- The buffer network and link data;
- Restricted turns and links;
- Node and zone co-ordinates;
- Bus routes; and,
- Generalised costs in the case of the multiple user classes.

Details of the network data file and how to program the information into the data file in the right format can be found in the SATURN User Manual (van Vliet and Hall, 1998). Within the restricted turns and links section links and/or turns may be banned from making certain movements or penalised to account for differences between HGVs and cars for instance.





**Figure 6.3: Map of the coarse DTO zones aggregated from the O-D trip demand matrix for the GDA transport network.**

The ability to be able to add restrictions to links is used to define the cordon location across the highway network. The restriction, a monetary or time penalty, is applied to

links that cross the cordon boundary in the same way that a real cordon scheme would charge at the boundary location. This defines the boundary location for the cordon scheme and adds the charge associated with that scheme to trips passing through the cordon when the Network.UFN file is combined with the trip demand matrix during modelling. An example of the coding used in the restricted turns/links section, and in this case, to define a cordon is also given in Section 10.2 of Appendix 10.

### 6.3.3.2 Model input: the trip demand matrix

The GDA network is made up of a number of zones. Each zone represents both an origin and a destination on the highway network. Originally, for the DTI Strategy, the GDA was divided up into 367 zones based on District Electoral Divisions (DEDs). A 432-zone model was then developed by the DTO for the 1997 base year Dublin transportation network model. The zone numbers, zone descriptions, the DEDs that were used to define the boundaries into which trip information from census data was gathered, and the CSO equivalent reference numbers are given in Appendix 11. The trip demand matrix is, in the case of the 1997 model, made up of  $432 \times 432$  trip elements representing the trips associated with each O-D pair, where each zone represents an origin and a destination to every other origin and destination in the matrix. The trip demand matrix zones can also be aggregated to produce a set of 21 coarse zones as seen in Figure 6.3. This Figure shows the area that is covered by the zone system. At the extremity of this map, matrix elements defining zones at the boundary of the GDA highway network are used to represent the trips entering and leaving the zoned area via Primary National Roads from other parts of the country. In other words, these routes are treated as external zones.

The trip demand matrix information is input to the parallel transport model in the form of an 'UnFormatted' Matrix (UFM) file, which will hence forth be known as the Matrix.UFM file. The Matrix.UFM file is produced using the Mx program, which is included in the SATURN Suite. Mx takes the information produced in a standard DAT file in a similar way to SATNET for network DAT files, and converts this text-formatted file to the Matrix.UFM file format, which is then input to the assignment and

simulation model. An example of one of the DAT files used to create a Matrix.UFM file can be found in Appendix 12.

In this research both HGVs and private cars are modelled on the highway network. This represents multiple user classes as each type of vehicle is called a class. In order to represent both the movements of private cars and HGVs, i.e., multiple user classes, two separate 432\*432 matrices are needed. These matrices are then stacked one on top of the other using the Mx program so that they can be used as input to the parallel transport model together, and hence, modelled across the highway network together.

A cordon application affects the trip demand; therefore, it affects the trip demand matrix. To express this in the trip demand matrix an elasticity demand function is used in conjunction with the original trip matrix, the costs matrix produced from the first all-or-nothing assignment defining the costs associated with each O-D trip, the charge associated with the cordon application, and, the price elasticity of user demand associated with that charge. This function is described in Section 6.3.5 below. The elasticity demand function is applied to the specific zones within the original trip matrix that refer to areas that are affected by the cordon application, i.e., zones inside the cordon ring. A set of new trip demands are produced for zones crossing the cordon boundary and this new set of values is then used to replace the existing trip demand values for the O-D pairs which have their destination inside the cordon boundary.

Due to the size of the trip matrices involved the HGV and private car trip matrices were broken down into manageable sections for manipulation as described and then reconstituted with the new trip demands included. Once this was achieved the new trip demand matrix, in standard txt format, was input to the program Mx to produce the Matrix.UFM file for input to the parallel model.

#### *6.3.4 Charge levels and the price elasticity of user response*

The main factors and constraints include the charge associated with the cordon or screen-line, the elasticity of user response to this level of charge and the elasticity

demand equation used to calculate the new trip demand resulting from the implementation of the congestion charging cordon scheme.

#### **6.3.4.1 Charge levels**

Road user charges should be linked as closely as possible to marginal external costs (Boot et al., 1999). A study of transport externalities in Dublin found that a charge of approximately €6.86 would be representative of the externalities associated with trips in the GDA (Gibbons, 1998). In London, consultants Halcrow Fox (1993), have found that between €8.20 – €16.39 (£5stg – £10stg) could be charged to enter central London. In Singapore, where they use a system with restricted zones and toll charges on the major expressways entering the city centre, the charges range from between €2.09 – €3.48 (S\$3.00 and S\$5.00) during peak periods (Willoughby, 2000). In 1999, Oscar Faber et al (1998) suggest charges varying from €0.79 – €3.81, but this report was based on prices in 1997. Further testing has also been carried out in Cambridge using charges between €0.98 – €8.20 (£0.60stg and £5.00stg) for congestion charging (May and Milne, 2000).

On the basis of the summary above, and on the review of road user charges associated with congestion charging carried out in Chapter 3, a relatively wide range of charges will be used for this research to reflect international experience. Charging levels have varied greatly across the case studies described in Section 3.6, from \$0.25 to £8stg. However, according to Santos (2000a), errors in estimating the effective charge will not hugely affect the efficiency of the system. Therefore, based on implemented examples of congestion charging, i.e., Singapore and London, a fixed charge in the range of €3 to €8 would be appropriate. This fixed charge will be applied only during the morning peak period from 8am to 9am due to the fact that the DTNM does not, at present, allow for the modelling of peak spreading or for a longer peak period.

#### **6.3.4.2 Price elasticity of user demand**

These elasticities are directly related to the overall generalised costs through the elasticity demand function described in the next section. During peak periods elasticities are lower because work related commuting trips, which make up 67% of

trips during this period (DTO, 1998), have to be made through one mode or another (Hall et al., 1992). 19% of trips are school or college related while other trips make up the remaining 14% (DTO, 1998). For a journey to work, there is little chance of a change in destination or abandonment of the trip, which suggests that the elasticity of demand for commuting with respect to changes in generalised cost may be relatively low (Oscar Faber et al, 1998). As such, disincentives, like a road user charge, are likely to have less effect on these high priority trips.

At this stage it is not possible to assign specific demand elasticity values to represent user response to specific congestion charges and specific congestion charging scenarios in the GDA. Given the need to make an assumption on elasticity values, they should be restricted to within a close range that is reasonably justified. Taking into account the literature and studies described in Section 3.7, for instance it is suggested in Oscar Faber et al. (1998) that a peak demand elasticity value of between  $-0.1$  and  $-0.5$  should be used for the GDA, as well as practical experience, a sensible range of values might be considered to be  $-0.2$  to  $-0.7$  for an inner city fixed price congestion charging scheme. This range of values will be used in conjunction with the elasticity demand function described in the next Section to represent trip maker's reactions to charges during the peak period between 8am and 9am on working days

### 6.3.5 *Elasticity demand function*

The elasticity demand function used for the research is the same as that used in Oscar Faber et al (1998). The elastic exponential, or semi-log function, referred to below is one of four demand response functions available as part of an elastic assignment algorithm in the SATURN Suite, see van Vliet (1982) and Hall et al. (1992).

The methodology used in the UK Engineering and Physical Sciences Research Council (EPSRC) study, that led Oscar Faber et al (1998) to use this function, involved building an independent relationship from a set of stated preference (SP) data via traditional logit-based approaches and then attempting to fit one of the four SATURN elastic assignment options to it. In the study, the semi-log function, described below, came out as a best fit. In a subsequent, more complex project where SP data was collected in

Cambridge, Norwich & York, the best fit was found to be a simple exponential function.

However, according to Oscar Faber et al (1998), the function below is more consistent with what might be expected in the case of the GDA and this is assumed to be the case for this research. The exponential aspect of the demand function also takes into account the fact that work, or priority, trips make up approximately 67% of the trips being made in the peak hour period (DTO, 1998). The function does this by dampening the effect that an increase in the elasticity of user response has on the trip demand to give a more realistic, or elastic, response to a cordon charge over a peak hour charging period.

$$T_{\text{new}} = T_{\text{base}} \text{EXP} (\text{El.} ( ( \text{CH} + C_{\text{base}}) / C_{\text{base}} - 1 ) ) \quad \text{Equation 6.1}$$

Where:

- $T_{\text{new}}$  is the trip demand after charging has been applied;
- $T_{\text{base}}$  is the trip demand in the base year per route;
- CH is the road user charge at the cordon;
- $C_{\text{base}}$  is the total generalised cost in the base year for each trip and; and,
- El. is the constant power coefficient that regulates the shape of the function, i.e., the price elasticity of user demand.

However, if considering small changes in demand the equation will not work well. This is because a small change in the charge level will result in a large change in the number of trips affected. This is due to the fact it is a logarithmic function.

### 6.3.6 Generalised cost

It is assumed that journeys are made for the utility (value) taken from activities that are performed at their destination. However, when making such a journey the traveller incurs the disutility of expense, time spent travelling, etc. Generalised cost is defined as a representation of this disutility. The zone-to-zone generalised cost used by the DTI is given by the two formulae (DTI, 1994):

$$GC_{car} = a_1IVT + a_2WK + a_3EX + a_4VOC + a_5PK + a_6 \quad \text{Equation 6.2}$$

$$GC_{pt} = a^1_1IVT + a^1_2WK + a^1_3WT + a^1_4FARE + a^1_6 \quad \text{Equation 6.3.}$$

Where:

- IVT is the in-vehicle travel time;
- WK is the walk time;
- WT is the waiting time for public transport;
- EX is the excess time while finding parking;
- VOC are the vehicle operating costs;
- FARE are the public transport fares;
- PK refers to the parking charges for the private car; and,
- $a_1$  to  $a_6$  are parameters used as both conversion factors and weights to express generalised costs in common units.

The generalised cost can be expressed in monetary or time units. If the generalised cost is measured in monetary terms 'a<sub>1</sub>' can be interpreted as the perceived value of IVT. Usually 'a<sub>2</sub>' and 'a<sub>3</sub>' are approximately twice that of 'a<sub>1</sub>', making the perceived value of walking and waiting times twice that of IVT. Parameters 'a<sub>4</sub>' and 'a<sub>5</sub>' convert values of fare into their equivalent values in time while 'a<sub>6</sub>' is the mode specific constant representing the inherent costs of a mode.

Due to the fact that the generalised cost is calculated as a function of time the congestion charges for crossing the cordon boundary are converted to time restrictions, in seconds, based on the values of time for the private car and HGV user. These values of time are based on those used by the DTO for the base year model definitions.

#### *6.3.7 Inner city congestion charging cordon testing*

The trip demand matrix file, Matrix.UFM, and the highway network file, Network.UFN, are used as inputs to the parallel implementation of the assignment/simulation section of the SATURN traffic model, SATALL, which allocates (assigns) trips to routes through the highway network.

The trip demand matrix and highway network files developed by the DTO are used for initial testing of the 1997, 2000 and 2006 year scenarios. The base year for these files is 1997 and the files were then factored up by the DTO to represent trip demand on the highway network for 2000 and 2006.

A number of changes have been made to the highway network file for the 2006 version of the GDA road network. The changes have been made to represent the completion of planned and confirmed infrastructure projects that are to be put in place in the intervening years. These changes include, amongst others, the completion of arterial and radial quality bus corridors (QBCs), improvements to suburban rail, completion of a cycle network and completion of two lines of the new light rail implementation that is already under construction. The DTO assumed that the growth for 2006 is 60% of the growth that occurs from 1997 to 2016. The 2016 values are based on projections, on a



zone-by-zone basis, in local authority areas and are capped at the Strategic Planning Guidelines totals (Brady Shipman Martin, 1999).

The testing procedure involves adapting both the Network.UFN file and Matrix.UFM file to accurately represent the different cordon charging scenarios applied to the GDA highway network, see the description of these processes in Section 6.3.3. The highway network file is edited to incorporate the cordon by applying a charge to links defining the cordon boundary, see Figure 6.1 and Figure 6.2 for the cordon location, while the trip matrix file is edited by applying the elasticity demand function, see Section 6.3.5, to the standard base trip matrix for the year in question to produce the new trip demand matrix for the cordoned network.

When this research began the DTO were not using elastic assignment to produce network flows and distributions. Consequently, there were no elastic parameters, or other necessary information, available to run elastic assignment for the DTNM. As a result, this research concentrated on the fixed assignment aspect of the SATURN traffic assignment and simulation model so that the new parallel model could be used to assess the DTNM. However, to incorporate an elastic response this research has applied the elasticity demand function described in Section 6.3.5 to the trip demand matrix. This allows for a more realistic response from users than having a linear response to the cordon charge.

This elasticity demand function is applied to the trip demand in the base year per route in conjunction with the road user charge associated with crossing the cordon, the price elasticity of user response and the total generalised cost in the base year per O-D pair. The generalised cost is found using the equations described in Section 6.3.6.

At the start of an elastic assignment or a fixed assignment procedure an all-or-nothing assignment is undertaken. This assigns trips to routes on the highway network in an all-or-nothing fashion. This results in total flow along links in the highway network and the costs associated with each link can then be calculated and then summed together to give the corresponding highway network costs for each O-D pair. This matrix of

generalised costs then represents the costs for each O-D route across the highway network.

In the elastic assignment within the SATALL program of the SATURN Suite, one of the four elastic demand functions given by the software is applied to the trip matrix using the new set of generalised costs, producing a new trip demand matrix and the process continues. Though the fixed assignment procedure was reprogrammed in parallel an attempt was made to reproduce elastic response manually by applying the elastic demand function described above to the trip demand matrix using the generalised costs matrix obtained from the first all-or-nothing assignment.

#### 6.3.8 *Test scenarios*

The testing that has been carried out is summarised in the 'Lists of tests' given in the following pages for the 2006, 2000 and 1997 Dublin transportation network model representations. All testing has been carried out to convergence and the inner city cordon application has been applied to the highway network using a number of different elasticity levels ranging from 0 to  $-0.7$ . The monetary congestion charge of between €3 and €8 is converted to the time equivalent so that all of the parameters in the generalised cost equations described in Section 6.3.6 are the same. The costs manipulated in the elasticity demand function are also expressed in units of time because the costs matrix, output from the first all-or-nothing assignment and used in conjunction with the elasticity demand function to produce the new trip demand matrix, is produced by the generalised cost equation.

Three charge levels were used as the fixed charge for crossing the cordon boundary: €3.80, €5.70 and €7.60. These charges represent a time restriction of 3000, 4500 and 6000 seconds, respectively, when converted to their equivalent values of time. These charge levels were then applied to the 2006, 2000 and 1997 base year transport model produced by the DTO in the elasticity range mentioned above. The results from these tests are reported in Section 6.4.

### 6.3.8.1 List of tests: 2006

A cross-reference list of the cordon tests carried out on the 2006 Base Year model produced by the DTO is given in Table 6.2. In addition to these tests the following results files were produced for the base case and for a test case:

- A06H-ds5: Base year trip matrix and network file for 2006; and,
- A06ic1: Test case adding a charge of €3.80 to the highway network definition and not applying any demand management reaction to the trip matrix.

Summaries of the results from these tests can be found in Appendix 9.

**Table 6.2: Details of the names of the test result files in the Appendix with reference to the road user charge and elasticity used.**

	No Charge	Ch = €3.80	Ch = €5.70	Ch = €7.60
No elasticity	A06H-ds5			
El. = -0.1		A06ic5	A06ic6	A06ic10
El. = -0.3		A06ic4	A06ic7	A06ic11
El. = -0.5		A06ic3	A06ic8	A06ic12
El. = -0.7		A06ic2	A06ic9	A06ic13

### 6.3.8.2 List of tests: 2000

A cross-reference list of the cordon tests carried out on the 2000 Base Year model produced by the DTO is given in Table 6.3. In addition to these tests the following results files were produced for the base year model without the cordon application and for a number of test cases:

- A00h-by0: Output from the base year trip matrix and network file for 2000, obtained from the DTO;
- A00ic4: Test case applying a charge of €3.80 and an elasticity of  $-0.5$  to the trip matrix but not to the network file to assess the effect, or lack of an effect on the model; and,
- A00ic5: Test case applying a charge of €3.80 to the network file, i.e., defining the cordon around the network and the penalty for crossing it, but not applying the effects of the charge to the trip matrix.

Summaries of the results from these tests can be found in Appendix 8.

**Table 6.3: Details of the names of the test result files in this Appendix with reference to the road user charge and elasticity used.**

	No Charge	Ch = €3.80	Ch = €5.70	Ch = €7.60
No elasticity	A00h-by0			
El. = -0.1		A00ic7	A00ic9	A00ic13
El. = -0.3		A00ic6	A00ic10	A00ic14
El. = -0.5		A00ic3	A00ic11	A00ic15
El. = -0.7		A00ic8	A00ic12	A00ic16

### 6.3.8.3 List of tests: 1997

A cross-reference list of the tests carried out on the 1997 Base Year model produced by the DTO is given in Table 6.4. In addition to these tests the following results files were produced for the base year model without the cordon application and for a test case:

- A97s182: Output from the base year trip matrix and network file for 1997, obtained from the DTO; and,

- A97h-by0: Test case applying a charge of 3000secs and an elasticity of  $-0.1$  to the trip matrix but not to the network file to assess the effect, or lack of an effect on the model.

Summaries of the results from these tests can be found in Appendix 7.

**Table 6.4: Details of the names of the test result files in this Appendix with reference to the road user charge and elasticity used.**

	No Charge	Ch = €3.80	Ch = €5.70	Ch = €7.60
No elasticity	A00h-by0			
El. = -0.1		A97ic0	A97ic4	A97ic8
El. = -0.3		A97ic1	A97ic5	A97ic9
El. = -0.5		A97ic2	A97ic6	A97ic10
El. = -0.7		A97ic3	A97ic7	A97ic11

## 6.4 Results and discussion

### 6.4.1 Introduction

The simulation time period is one hour in duration, representing the morning peak period from 8am to 9am. The model continues to run until over-capacity queues at intersections dissipate completely, i.e., that the traffic on the highway network is running without over-capacity queuing on roads between junctions (links) or junctions (nodes). However, from discussions with the DTO and Dr. D. van Vliet it was recommend that results obtained from this period not be taken into context. Therefore, the results described here will concentrate on the simulation period only.

The road user pricing scenarios will be assessed and discussed using a number of variables described here:

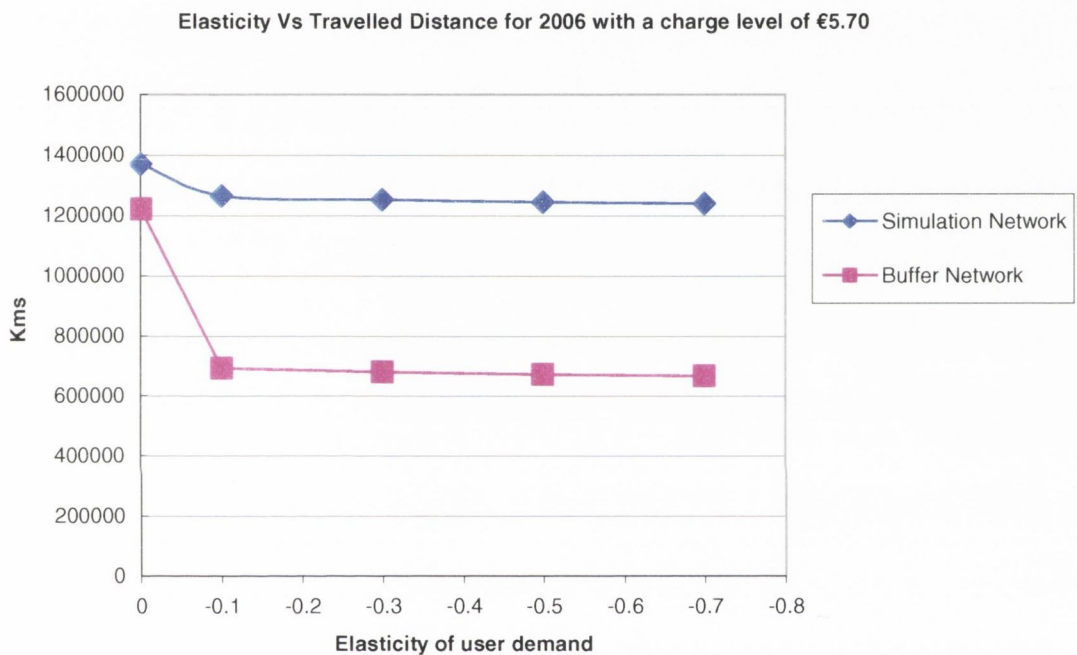
- Transient queues record the time spent by vehicles in queues which, in the case of traffic signals, clear during a single cycle;
- Over-capacity queues record the extra time spent in queues at over-capacity junctions waiting for the cycle in which the vehicle exits;
- Free-flow time is the time which would be spent travelling on links operating at their free-flow speeds;
- Delays, the flow-specific extra travel time on those links with link speed-flow curves;
- The link cruise time is the sum of the free-flow time and the flow-specific extra travel time on those links with speed-flow curves;
- The total travel time is the sum of both link and junction times;
- The travel distance is the vehicle or pcu.kms on simulation links; and,
- The overall average speed is defined by the travel distance divided by the total travel time.

This results section will predominantly assess the effect of the inner city cordon application on the 2006 highway network and trip demand matrix definitions provided by the DTO. Testing has also been carried out on both the base year 1997 and 2000 trip matrix and highway definitions during model testing. However, since similar reactions have occurred over the DTNM for the 1997 and 2000 instances, and since the 2006 model specification is the most recent, the 2006 model will be discussed in detail in the following sections and the results from modelling the 1997 and 2000 year models are given in summary format in Appendix 7 and Appendix 8, respectively.

#### 6.4.2 Distances travelled over the highway network

Figure 6.4 shows reductions in the distances travelled over the simulation network and over the buffer network for a cordon charge of €5.70, with the elasticity of user

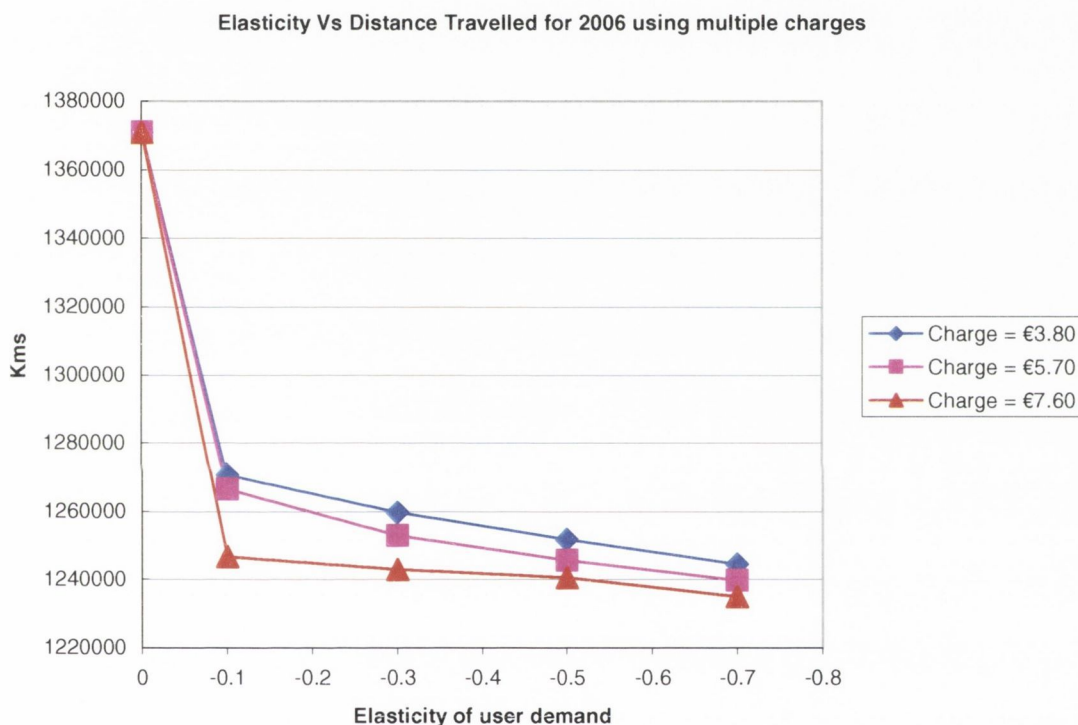
response to this charge ranging from  $-0.1$  to  $-0.7$ , respectively. The graph shows that the travelled distance over the buffer network drops sharply with the inclusion of the inner city cordon. This represents a 43% reduction in the distances travelled in the buffer network when compared to the standard model predictions without any cordon application.



**Figure 6.4: The effect of elasticity on the distance travelled during the simulation period for the simulation network and the buffer network.**

The simulation also shows a reduction of 8% for an elasticity of  $-0.1$ , increasing to 10% when an elasticity of  $-0.7$  is applied. The simulation network includes the area inside the M50 Motorway, shown in Figure 6.1 and in Figure 6.2. The change from the simulation network to the buffer network takes place here. Therefore, most of the links entering the simulation area are main radial routes that terminate in the city centre. As such, they are heavily affected by the inner city cordon application. This appears to represent the fact that less trip makers are making the trip to the city centre by car and HGV from the buffer network. In comparison to this the distances travelled in the simulation network drops by 8%, or approximately 130,000 Kms. It is noted from

Figure 6.1 that the simulation area is substantially larger than the cordoned area. It is most likely, therefore, that trip makers wishing to avoid the cordon charge are re-routing around the cordon while remaining inside the simulation network, travelling further to get to their destination. This leads to increased travel along orbital routes in the simulation network.



**Figure 6.5:** The effect of elasticity on the distances travelled over the simulation network using charges of €3.80, €5.70 and €7.60.

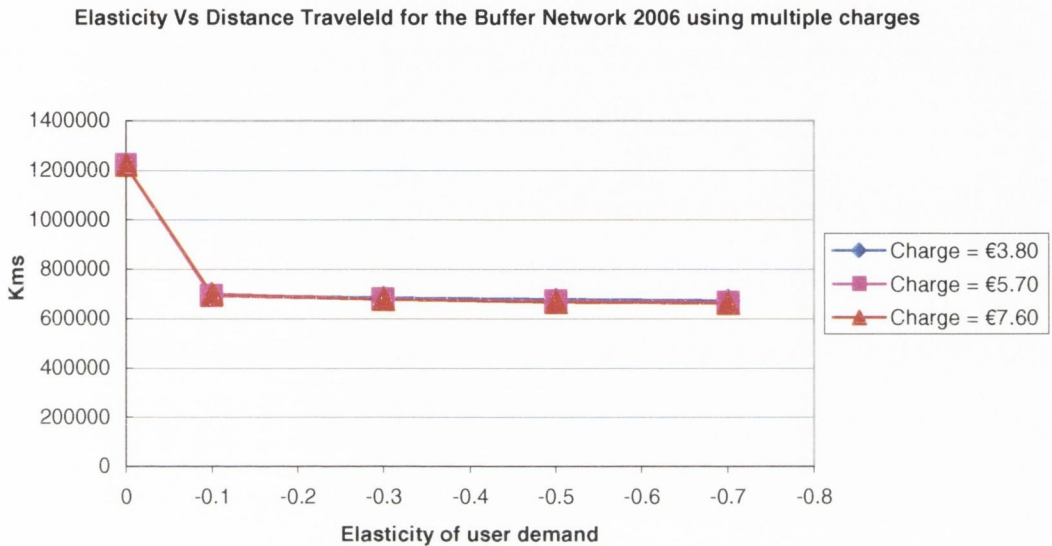


Figure 6.5 shows the reduction in travelled distances as a result of the cordon application for the simulation network using three different charges; €3.80, €5.70 and €7.60. As expected, the higher the charge, the more effect the cordon has on the distances travelled. It is noted, however, that the differences in the effectiveness of the cordon application across the three different charges reduces as the elasticity is increased. This reflects the nature of the trip makers during the peak period. Work related trips are considered to have a high priority, higher than leisure for instance. As such, a certain high percentage of trips have to be made, the only variable open to interpretation for these trip makers is what route to take.

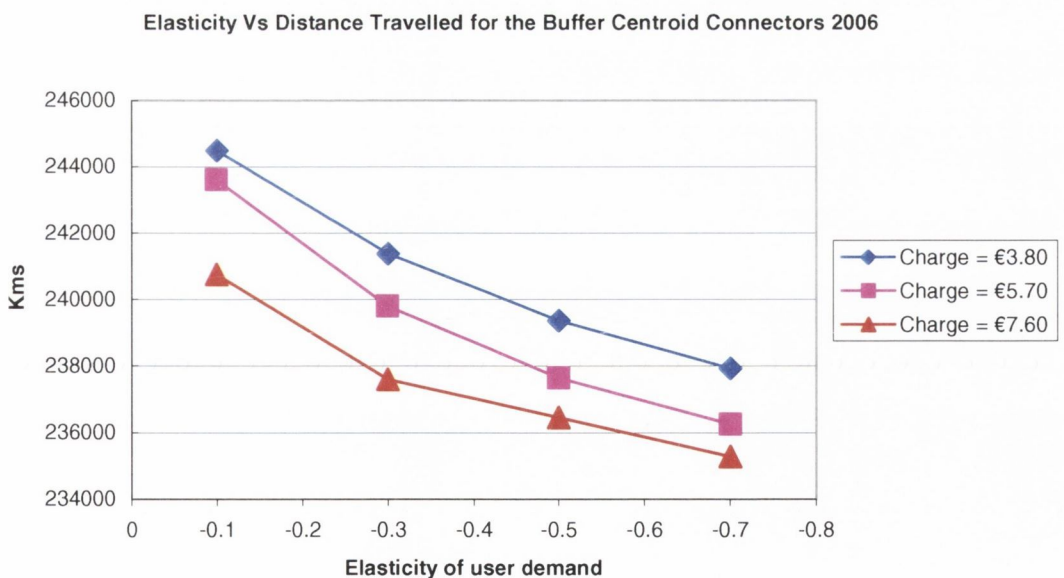
Figure 6.5 also shows an initial drop in the distance travelled for each of the three charges using an elasticity of  $-0.1$ . This is expected in that any cordon implementation will have an initial reaction, especially from trip makers who do not have to make a priority trip. However, there are only a certain number of trips that are affected by the inner city cordon. As a result, once trip makers who will not, or can not, pay the cordon charge are priced off the network, i.e., either using another mode of transport, cancelling their trip or making the trip at another time, the decrease in the distances travelled resulting from further trip reductions levels out quickly as the majority of trip makers have a high priority trip to make or are unaffected by the cordon.

Figure 6.6 shows the effects of the same charge levels on the buffer network. It is immediately apparent that the travelled distances on the buffer network vary little in terms of the charge used or in terms of the elasticity applied. While there is a slight reduction in the travelled distances for further elasticities and for the charges used, in general, the differences are small. This can be explained with the aid of Figure 6.7 and Table 6.5. In the case of the buffer network nodes are not modelled. Only links and link related delays are modelled. The distance travelled in the buffer network remains at relatively the same level because trip makers are queuing to get onto the network itself as opposed to queuing on it. This occurs, in particular, at the exits to large housing estates where large queues form to get onto the highway network. Figure 6.7 shows the effect of the elasticity of user demand on the distance travelled across the buffer centroid connectors. Centroid connectors connect zones to the highway network

and are used to represent this queuing to get onto the buffer network. As can be seen from Figure 6.7 and Table 6.5, which gives the numerical values for this graph,



**Figure 6.6:** The effect of the elasticity on the distance travelled over the buffer network using charges of €3.80, €5.70 and €7.60.



**Figure 6.7:** The effect of the elasticity on the distance travelled over the buffer centroid connectors using charges of €3.80, €5.70 and €7.60.

**Table 6.5: Table of the distances travelled over the buffer centroid connectors using charges of €3.80, €5.70 and €7.60.**

Elasticity	Distance Travelled over Buffer Centroid Connectors, in Kms		
	Charge = € 3.80	Charge = € 5.70	Charge = € 7.60
-0.1	244498	243606	240763
-0.3	241382	239794	237597
-0.5	239372	237635	236462
-0.7	237924	236260	235286

the distances travelled along these connectors are dropping as the elasticity increases and as the charge levied at the cordon increases. The reason that the effects are not more prominent in Figure 6.7 is simply because most of the trip makers are not travelling to the city centre and those that are have already responded more to the initial introduction of the cordon than to the subsequent increases in charge.

Overall, the simulation network has approximately an 8-10% reduction in the kilometres travelled for the 2006 transport model with an inner city cordon scheme with the charge level ranging from €3.80 to €7.60, respectively. Due to the fact that the simulation area is so much larger than the cordoned area this reduction seems to suggest that a cordon scheme would be very effective and implies an even higher reduction inside the cordon boundary.

### 6.4.3 Travel time

The total travel time on the highway network is measured in PCU.HRs. PCU stands for Passenger Car Unit, a standard unit for the measurement of traffic levels. It is recognized internationally that the presence of HGVs and slow-moving vehicles greatly influences the design of a highway. Therefore, in this modelling process, the results are

reported using PCUs. The flow coefficients shown in Table 6.6 are used to convert vehicles into PCUs.

The total travel time for the simulation network is given in Figure 6.8 using the cordon charge levels of €3.80, €5.70 and €7.60 and elasticities ranging from  $-0.1$  to  $-0.7$  on the 2006 transport network model. It is apparent from Figure 6.8 that the total travel time decreases as expected with the initial application of the cordon model but continues to fall steadily as the charge or elasticity is increased. This is a similar response to that found for the distances travelled over the simulation network. The combination of the reduction in the distances travelled and the total time spent on the network during the peak hour suggests that either the highway network has become so over-congested and over-capacity that the traffic has come to a halt, or, that there are less trips on the network from private cars and HGVs as a result of the cordon application. It is assumed that the latter of these cases is most likely though an assessment of the average speeds over the highway network will help to determine the validity of this assumption, see the next Section on speed.

**Table 6.6: Flow coefficients for PCU measurement.**

Vehicle type	Flow coefficient (PCU)
Bicycles	0.5
Motorcycles	0.5
Light, commercial motor vehicles (gross weight 10 tons)	1
Passenger cars	1
Trucks and buses	2
Semi-trailers and trailers	3

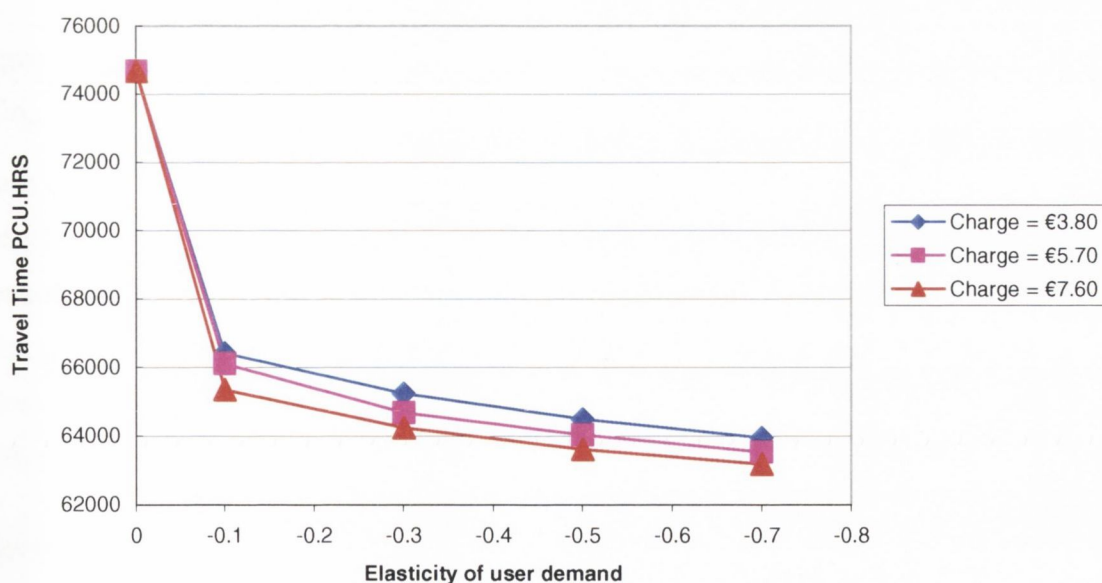
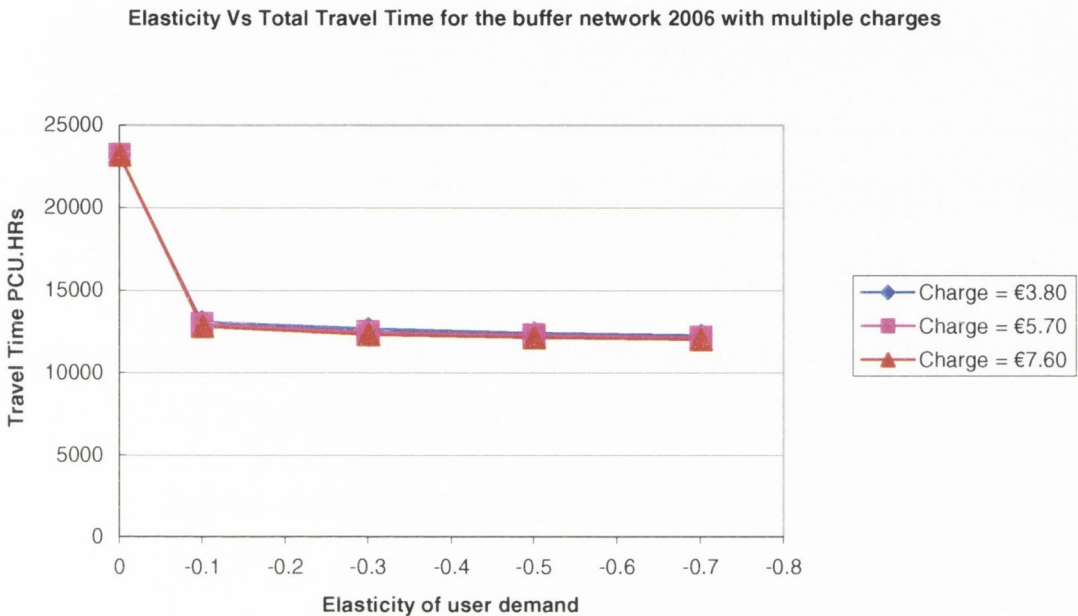
**Elasticity Vs Total Travel Time for the simulation network 2006 with multiple charges****Figure 6.8: The effect of the elasticity on the total travel time in the simulation network using charges of €3.80, €5.70 and €7.60.**

Figure 6.9 gives the total travel times for the buffer network and these times show a predictable outcome with little overall reduction in the total time travelled on the buffer

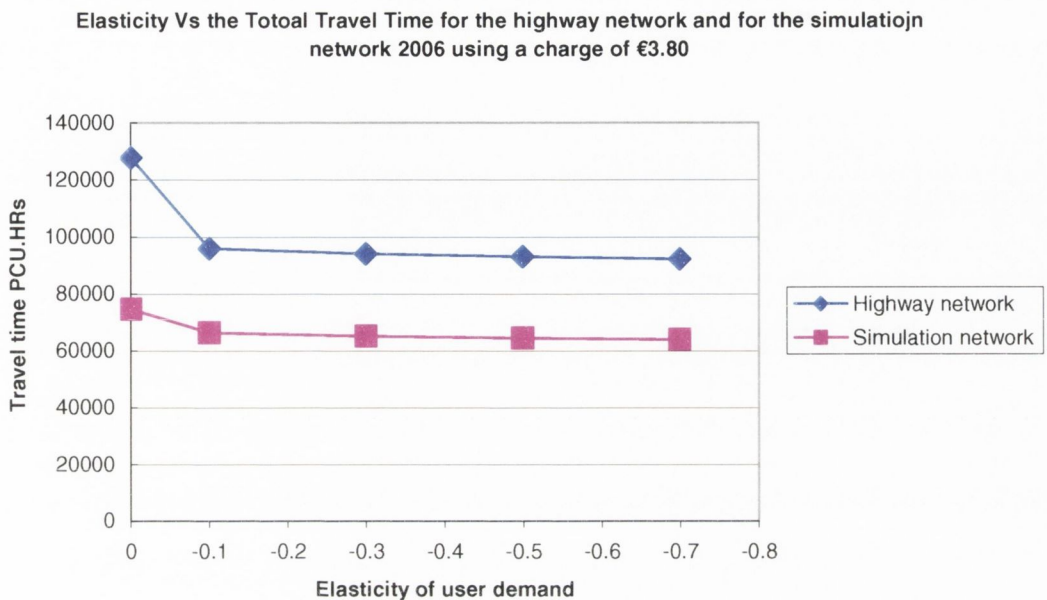
network after the initial implementation of the cordon because trip makers are waiting to get onto the network from buffer centroid connectors as discussed previously.



**Figure 6.9: The effect of the elasticity on the total travel time in the buffer network using charges of €3.80, €5.70 and €7.60.**

During the simulated peak period there is a 15-16% reduction in the total travel time in the simulation network for a charge of €5.70 implying a larger reduction within the boundaries of the cordon. This compares favourably to results obtained in London where congestion charging is predicted to cut traffic levels inside the charging zone, in kilometres travelled, by 10-15% and congestion, measured in 'vehicle delays', by 20-30% (GOL, 2000).

Figure 6.10 shows the overall picture of the reductions in total travel demand using a cordon charge of €3.80 on the 2006 highway network. This demonstrates an overall reduction in the travel time on the highway network of approximately 27% for an elasticity of  $-0.7$  and 25% for an elasticity of  $-0.3$ .



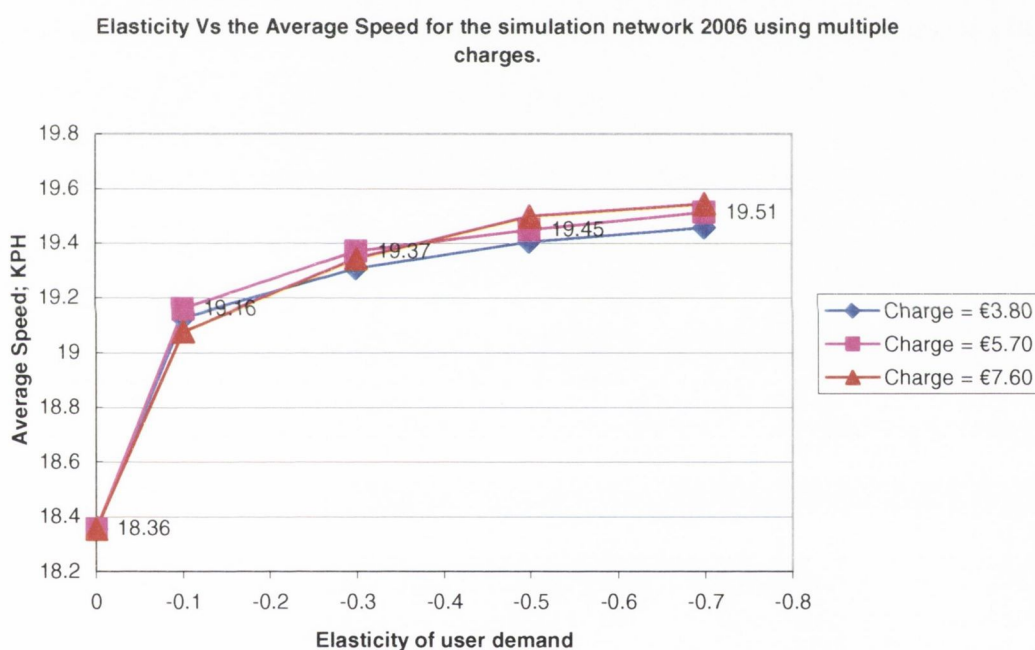
**Figure 6.10: The effect of efficiency on the total travelled time for the complete highway network and for the simulation network using a charge of €3.80.**

#### 6.4.4 Speed over the highway network

The average speed is simply calculated by dividing the total distance travelled by the total travel time giving KPH, (kilometres per hour). Figure 6.11 gives the average speed for the 2006 simulation network over the elasticities and charges described previously. This graph shows that the average speed increases steadily as the elasticity is increased. This is true for any of the charges used, and indeed, the increases in charge levels from €3.80 to €7.60 appear to have relatively less effect on the average speed in comparison to the elasticity. Santos (2000) makes reference to the fact that errors in choosing the correct charge are unlikely to influence the outcome of modelling tests to any great extent. This certainly appears to be the case here. The fact that the speed is increasing in the simulation network indicates that traffic congestion is reducing as trip makers are, relatively speaking, travelling further in less time. An increase in the average speed of just over 1 KPH is felt for a charge of €5.70 and an elasticity of  $-0.3$ . If this value of average speed is to be increased then it would appear that either a larger inner city cordon, a second cordon located at, or close to, the only other geographically suitable location at the M50 Motorway, or the use of screen lines

jutting out from the inner city cordon location, used for this research, to catch more orbital trip makers trying to avoid the cordon charge.

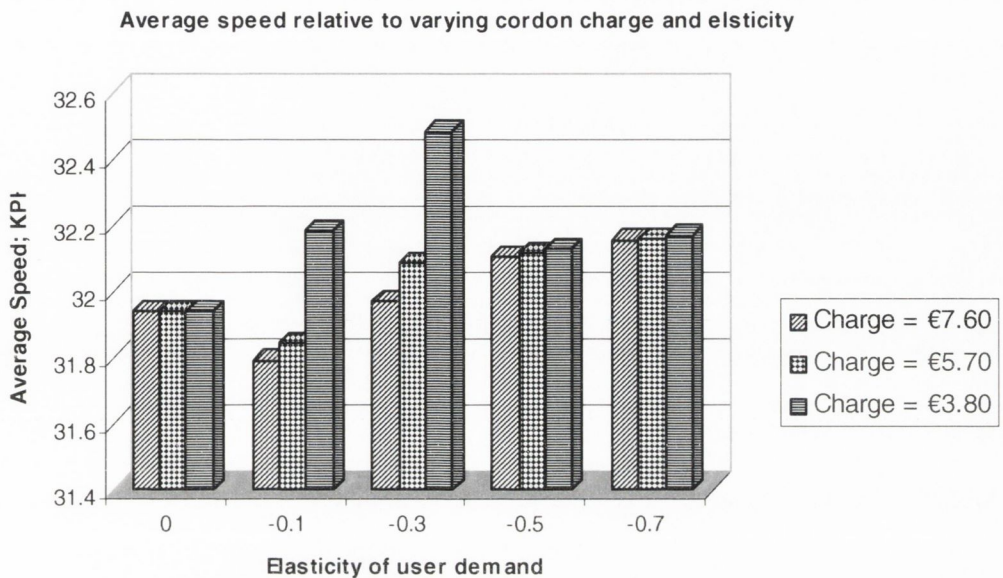
Figure 6.12 describes the effects of the cordon application on the buffer network, including delays experienced on buffer centroid connectors. The three charge levels used for this modelling are applied to the 2006 model for each elasticity increase modelled. For instance, for an elasticity of 0, there is no change in the average speed for each of the three charges because they do not apply. As the elasticity is increased relative to the three charges the average speed varies from just below



**Figure 6.11: The effect of elasticity and charge on the average speed of the 2006 traffic network.**

32 KPH to just above it. These fluctuations in the average speed, showing only small improvements, reinforce the conclusion that during the simulated peak period large queues form on centroid connectors that are unaffected by the cordon application. Therefore, due to the size of these queues, the relative effect of increases in the buffer network average speed, see Figure 6.13, due to the cordon application are largely negated by trip makers queuing to try to get onto the buffer network itself.

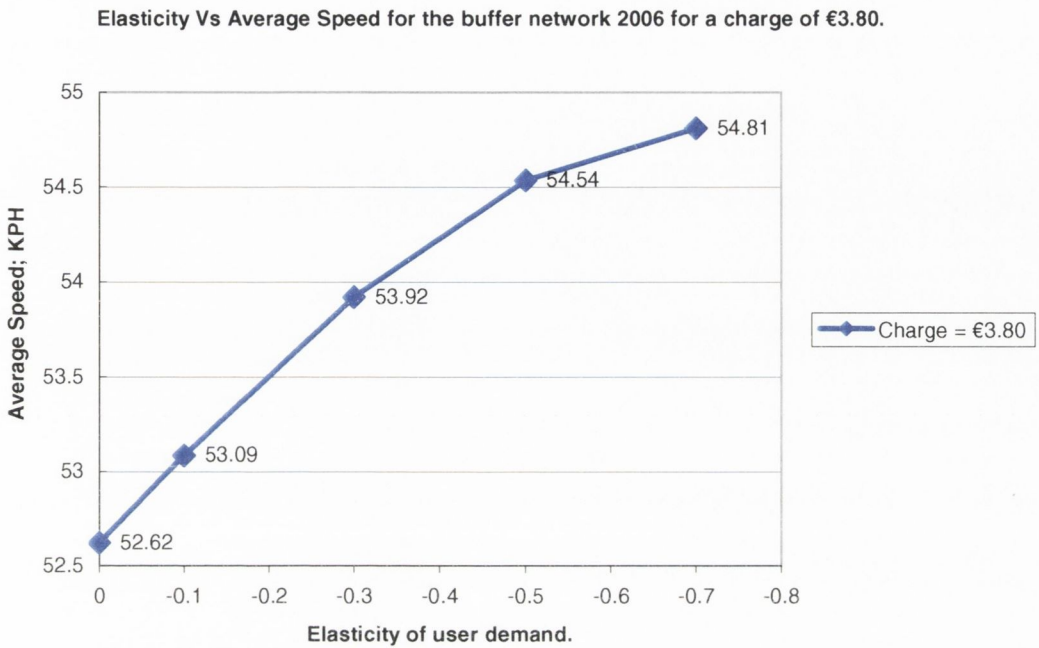




**Figure 6.12: The average speed on the buffer network 2006, including delays experienced on buffer centroid connectors, using charges of €3.80, €5.70 and €7.60.**

An average speed increase of over 2 KPH is felt on the buffer network when considered without centroid connectors. This compares to a maximum increase shown in Figure 6.12 of just 0.5 KPH. This increase in the average speed of 2 KPH for traffic flows is found using a cordon charge of €3.80 and an elasticity of between  $-0.5$  and  $-0.7$ . Centroid connectors are not included in order to examine the buffer network without the delays associated with buffer centroid connectors during the peak hour, which have the effect of reducing the average speed.

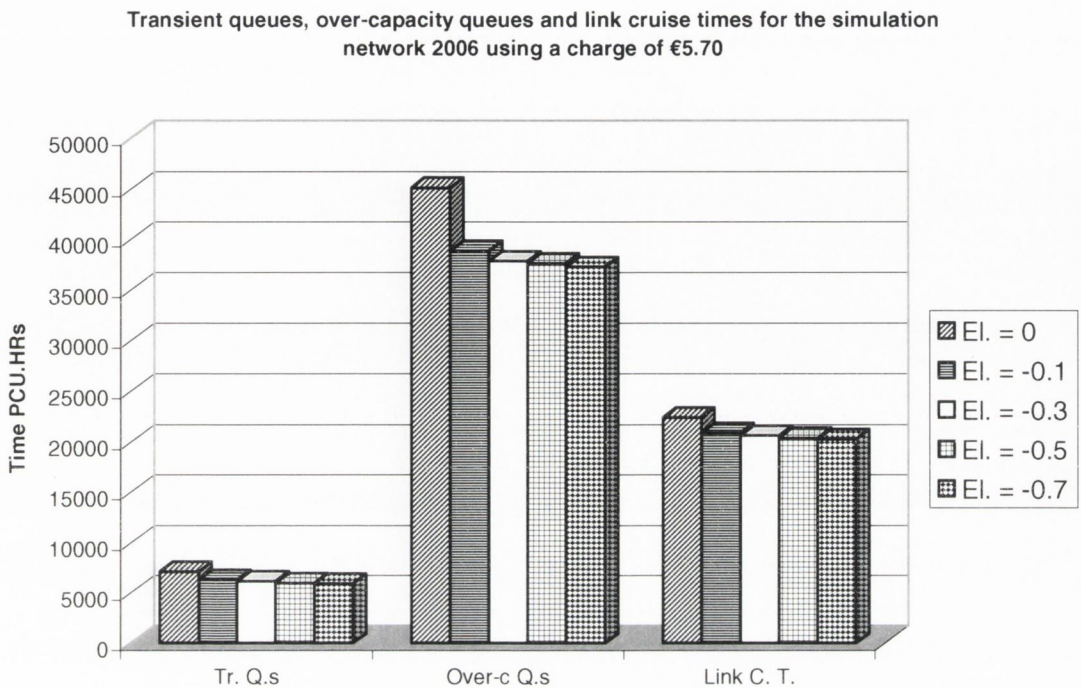
Figure 6.11 and Figure 6.12 also demonstrate that for the average speed, the elasticity of user response or the charge level do not have a significant effect relative to the affect that the demand function itself has.



**Figure 6.13:** The average speed across the buffer network excluding buffer centroid connectors using a charge of €3.80.

#### 6.4.5 *Queuing*

There are three types of queues that are examined in this Section. There are transient queues, over-capacity queues, and, link cruise times, which are a combination of the free flow times along links and the delays encountered along those links. See Section 6.4.1 for a more thorough definition. Figure 6.14 shows the transient queues, over-capacity queues and link cruise times, in parts 1, 2 and 3 respectively, from a cordon application of €5.70 on the 2006 simulation network.

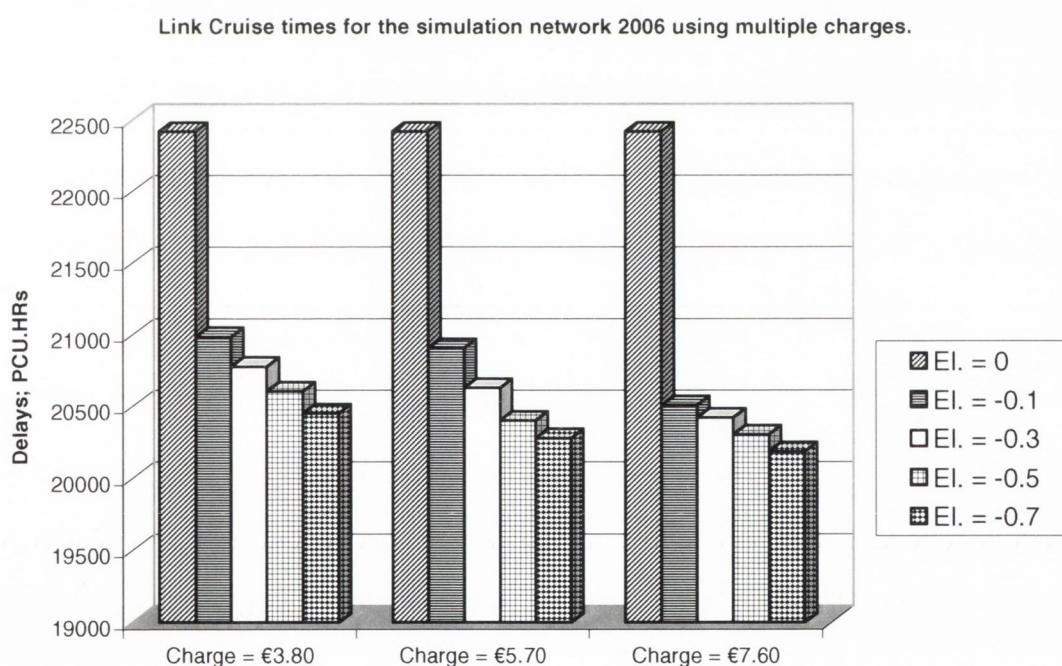


**Figure 6.14: Graph of the transient queues, over-capacity queues and link cruise times associated with the application of a €5.70 cordon to the 2006 highway network.**

It can be seen from Figure 6.14 that the elasticity demand function is reducing the effect of higher elasticities to represent the response of trip makers to the cordon charge during the morning peak period. Part 1 of Figure 6.14 shows that transient queues remain at relatively the same level irrespective of increasing elasticity. This is due to the fact that the morning peak period in the GDA lasts longer than the peak hour modelled from 8am to 9am. Therefore, transient queues, representing the time spent by vehicles in queues that, in the case of traffic signals, clear during a single cycle, remain high as there is a constant demand for road space and the highway network remains close to capacity beyond the one hour peak period modelled. The very slight reductions in transient queues signify that some junctions are beginning to operate at less than capacity. This is discernable because transient queues are the last queues to dissipate on the network after over-capacity queuing. In other words, there must be less cars queuing at a particular junction than can get through that junction in the cycle time

assigned to it for the transient queue for that junction to reduce. This appears to be happening on the simulation network as other queues dissipate.

From Figure 6.14 it is noted that there is little change in the link cruise times for the simulation network during the peak hour modelled. Figure 6.15 looks at the link cruise times over the three charge levels assessed. It is noted that the cordon application results in an initial drop as expected, which is far smaller than it appears in this graph, see Figure 6.14, part 3 or Figure 6.15. This is followed by stepped reductions as the elasticity is increased. The levels of charge, however, do not appear to have a lot of impact on the link cruise times in comparison to the other parameters. This leads to the conclusion that while the demand function itself has the most affect on the results, within the function it is the elasticity of user response that affects the number of trip makers more than changes in the charge.

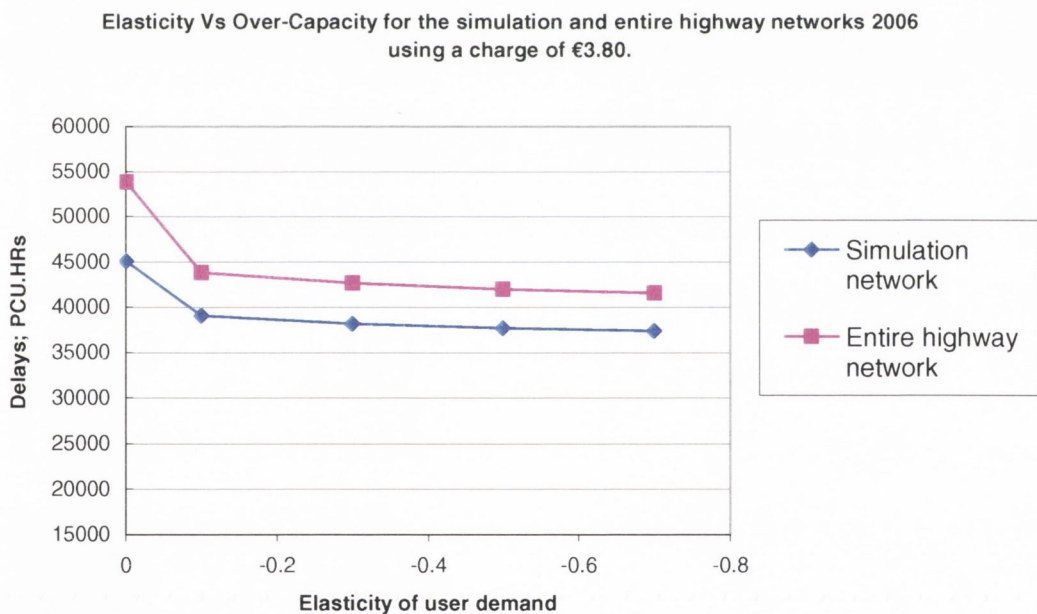


**Figure 6.15: Link cruise times for the simulation network 2006 using charges of €3.80, €5.70 and €7.60.**

Part 2 of Figure 6.14 gives the over-capacity queue totals for the simulation network. These queues represent over capacity links and junctions on the highway network

during the peak period and, as such, give a good indication of congestion on the network. In the case shown in Figure 16.4, applying the cordon with a €5.70 charge, there is approximately a 17% reduction in the over-capacity queues for an elasticity of  $-0.7$  and a reduction of approximately 14% when an elasticity of  $-0.3$  is used. This reduction takes place in the simulation network, which can be seen more clearly in Figure 6.16.

Figure 6.16 compares the over-capacity queue reductions from the entire highway network and from the simulation network. Therefore, as the totals from the simulation network go forward to making up the totals for the entire highway network it is apparent that the major reduction in over-capacity takes place in the simulation network.



**Figure 6.16: The effect of elasticity on the over-capacity queuing on the simulation and entire highway networks using a cordon charge of €3.80.**

#### 6.4.6 Fuel consumption

Fuel consumption is estimated for the simulation network using the following linear equation provided by the transport model:

$$TFC = FLPK*TTD + FLPH*TDT + FLPPS*S1 + FLPSS*S2 \quad \text{Equation 6.4}$$

Where:

- TFC is equal to the fuel consumption in litres per hour;
- TTD is equal to the total travel distance per hour;
- FLPK is equal to 0.07 litres/Km;
- TDT is equal to the total delayed time (the time spent stationary in queues) per hour;
- FLPH is equal to 1.20 litres/HR;
- S1 is equal to the number of primary stops at junctions;
- FLPPS is equal to 0.016 litres per primary stop;
- S2 is equal to the number of secondary stops at junctions; and,
- FLPSS is equal to 0.005 litres per 'second stop'.

Table 6.7 describes the fuel consumption for the three charge levels tested. Drops in fuel consumption of 12.5%, 13.1% and 13.5%, respectively, were achieved as a result of the cordon application. It is recognised that this estimate of fuel consumption is a crude one and is most likely out of date; van Vliet (1998) makes reference to Ferreira (1981) for more details on the parameters used in the equation. This compares when compared to other studies, for instance, in Edinburgh a reduction in environmental pollutants of between 3.5% and 8.3% has been predicted for a city centre cordon scheme with a charge of £1stg and £4stg, respectively (Edinburgh City Council, 2001).

Table 6.7: Fuel consumption statistics for 2006.

Simulation network	Charge = € 3.70	Charge = € 5.70	Charge = € 7.60
<b>Elasticity</b>			
0	180658.5	180658.5	180658.5
-0.1	162966.6	162269.3	159723.1
-0.3	160854.2	159614.4	158291.6
-0.5	159379.7	157879.1	157047.1
-0.7	158072.2	157002.2	156311.8

## *Chapter 7*

### *Conclusions*



### *7.1 Conclusions from parallel implementation*

The parallel transport network model is converging accurately, as expected, and with at least the same measure of accuracy that the sequential SATURN model is able to achieve. However, it also appears that as more processors are used the rate of convergence increases slightly by 1% or 2%, though the difference in convergence performance is not particularly significant.

The output from the parallel transport network model is within 1% of the results obtained from the sequential model for the scenarios tested over a number of runs. It is not clear why there is a difference at all but this, in itself, is expected because of the small differences in the convergence rates. It is noted that slight differences in results are obtained when the sequential model is run using the same input data, which implies that the differences are inherent in the sequential model itself and not because of the parallel implementation process.

The speed-up of the parallel transport network model (combining the sequential section and the parallel section) when compared with the number of processors being used is approximately equal to 5 for the program running over 16 processors. This represents an 80% reduction in the time taken for a single processor to complete the model run.

When a similar comparison is undertaken examining the parallelised module only, i.e., excluding the sequential section of the model, which is usually the case for performance analysis of parallel implementations, an increase in speed-up to approximately 8.16 is found for 16 processors. This compares favourably with results from a similar parallel implementation on a different assignment model and with the initial speed-up estimate for the parallel implementation, which was approximately 7.

The relationship between the efficiency and execution time for the parallel transport network model gives a ratio of 0.0129 when the parallel model is run over 16 processors. Due to the fact the efficiency – execution time ratio,  $R$ , for the parallel model run over 8 processors is 0.0127, a difference of just 0.0002, a tentative conclusion is reached that the ‘knee’, or average parallelism, of the parallel model lies

close to 0.013. Also, given the rate of conversion of this ratio for the parallel model, it is concluded that the ‘knee’ would most likely be found for the model tested using approximately 24 to 32 processors, which represents an opportunity for further research.

For the parallel section, the ratio is increasing with additional processors leading to the conclusion that the number of processors needed to reach the ‘knee’ for the parallel section is far more than is available for this research. As such it is hard to establish the effect of communication overheads on the parallel section of the model without actually applying the model in practice to more processors. Future work could, therefore, concentrate on finding the average parallelism, or ‘knee’, of the parallel section of the model and reducing the time spent in other sequential areas of the parallel model.

With respect to shortest path theory, it is concluded from the literature review of parallelisation within the field, see Section 4.7, that any attempt to parallelise the shortest path algorithm itself is unlikely to prove worthwhile from a performance standpoint.

Finally, the most computationally intensive part of the SATURN code, the SATALL program, has been successfully reprogrammed in parallel and can be ported between the majority of existing parallel systems. However, part of the parallel model remains in sequential format and this sequential section enforces harsh penalties on the performance of the parallel implementation. However, further work could include reprogramming of the remaining sequential sections to run in parallel. These sections include, for example, issues associated with I/O, the essentially sequential nature of the overall assignment procedure itself and the simulation section of the model.

## ***7.2 Conclusion from inner city congestion charging cordon implementation***

Based on the results of this research it is suggested that demand management could constitute an important part of an integrated solution for the congestion difficulties in the greater Dublin area (GDA). This is reinforced by the results of the tests discussed in Chapter 6 of this research.

There is a reduction in the number of trips entering the simulation network, as expected. However, because transient queue levels drop off slowly it is evident that trip demand within the GDA is remaining high. Considering that the peak period from 8 am to 9am is being modelled this is as expected.

There is a reduction of congestion delays on the simulation network during the charged period. This is demonstrated by the reduction in over-capacity queues on the simulation network of 14% for a charge of €5.70 and an elasticity of  $-0.3$ . Very slight reductions in transient queues signify that some junctions are beginning to operate at less than capacity. It is clear that the major reduction in over-capacity queuing taking place in the simulation network represents a reduction in delays at junctions and along the highway network.

A reduction of travellers is felt more quickly on the buffer network than on the more congested simulation network where there is an increase in the travelled distance between the cordon and simulation network boundaries. There is a 43% reduction in the distances travelled in the buffer network when compared to the standard model predictions without any cordon application.

The travelled distance, which also reduces slowly after an initial drop, demonstrates that people are travelling out of their way to avoid the cordon charge. However, because of the effects within the cordoned area, traffic inside the simulation area as a whole is moving more efficiently during the charged period as is demonstrated by the increase in the average speed from 18.4-19.5km/hr. It is also found from a comparison of the increases in average speed over the range of elasticities and charge levels used that the cordon charge has less affect on the highway network than changes in the elasticity of user response itself. However, it is the demand function itself that has the most effect on the overall results. Therefore, further research on different types of user demand function would be appropriate to assess the most realistic road user responses.

There are a substantial number of trip makers that are circumnavigating the cordon to avoid the charge. This is implied from the relatively low increase in the average speed compared to the comparatively larger reductions in over-capacity queuing. Some form

of radial screen-lines from the city centre cordon projecting outwards or an additional loop cordon application situated further out from the inner city may counteract these effects, which should be investigated further.

The buffer network appears to gain a lot of benefit from the inner city cordon and during the charged period the region inside the cordon also benefits greatly even though demand remains high. While there is a reduction of 8% in the travelled distance in the simulation network there is a 24% reduction across the complete network for a charge of €5.70 and an elasticity of  $-0.3$ . This compares very well with that experienced in London where very similar results were obtained. From modelling carried out a 20% reduction in car trips was modelled for Central London with a 10-12% reduction in total traffic level. In practice London's city centre cordon scheme was even more effective with a 16% reduction of traffic levels and a 32% reduction in overall congestion for an inner city £5 congestion charge. However, a 24% reduction for the GDA could be considered to be relatively accurate when the conclusions reached by O'Mahony et al. (2000), where a 22% reduction in kilometres travelled was found using a field trial conducted within the GDA, are taken into account. Similarly, results from a field trial in Leicester (Smith and Burton, 1998) also found reductions of approximately 18-22% for a morning peak period congestion charge.

An overall reduction in the travel time on the highway network of approximately 27% for an elasticity of  $-0.7$  and 25% for an elasticity of  $-0.3$  is felt when a charge of €3.80 is applied. However, the relative effect of increases in the buffer network average speed due to the cordon application are largely negated by trip makers queuing to try to get onto the network itself. This signifies that at a local level in the buffer network the effects of the cordon application are not felt as directly as elsewhere and large localised tailbacks are still prominent, particularly trying to get out of large housing estates in the morning peak period.

The analysis of fuel consumption suggests a reduction of approximately 12.5-13.5%, which is high compared to other studies of emissions. The model equation used is, however, only a rough estimate according to Dr. van Vliet and any serious conclusions

about emission rates would have to incorporate further research. This research could include an assessment of the likely impact of reductions of Sulphur Dioxide (SO<sub>2</sub>) and Nitrogen Oxide (NO<sub>x</sub>) in the atmosphere as a result of the cordon application.

Overall, the simulation network has approximately an 8-10% reduction in the kilometres travelled and a reduction of 15-16% in the total travelled time over the network for the 2006 transport model with an inner city cordon scheme using a charge of €5.70 and elasticities ranging from -0.1 to -0.7. These results appear to be relatively accurate based on previous studies and practical experience. For instance, in Dublin Oscar Faber et al. (1998) predicted a 21% vehicle hour reduction in the metropolitan region of the GDA using an outer cordon and a 12-25% reduction using an inner cordon. Comparable figures were also found in London with a reduction of 16% in Central London. Similarly, for modelling in Leeds, a city centre reduction of 11% was achieved for a peak period charge.

Considering this reduction in congestion it is noted that the simulation area is much larger than the cordoned area, which seems to suggest that a cordon scheme would be very effective and implies that there is an even higher reduction within the cordon boundary, as has been achieved for the Central London scheme where an overall congestion reduction of 32% was found, though this needs to be investigated further.

In addition to this, an assessment of the likely returns in revenue from such a cordon application would also be insightful as would research on mode split and the effects on other forms of transport such as public transport, cycling and walking.

## *References*

## References

- Adamson, P. and E. Tick (1991), Greedy partitioned algorithms for the shortest path problem, *International Journal of Parallel Programming*, Vol. 20, pp. 271-298.
- Ahlstrand, I. (2001), The politics and economics of transport investment and pricing in Stockholm, *Journal of Transport Economics and Policy*, Vol. 32, Issue 3, pp. 437-485.
- Ahuja, R. K., Magnanti, T. L., and J. B. Orlin (1993), *Network Flows, Theory, Algorithms and Applications*, Prentice-Hall.
- Allsop, R. E. (1997), Guest Editorial, *Transportation Research C*, Vol. 5, No. 2, pp. 73-76.
- Arentze, T., Hofman, F. and H. Timmermans (2003), *Predicting Multi-Faceted Activity-Travel Adjustment Strategies in Response to Possible Congestion Pricing Scenarios using an Internet-Based Stated Adaptation Experiment*, TRB 2003 Annual Meeting CD-ROM.
- Arezki Y. and D. van Vliet (1990), A full analytical implementation of the PARTAN/Frank-Wolfe algorithm for equilibrium assignment, *Transportation Science*, Vol. 24, No. 1, pp. 58-62.
- Baron, J. (1995), Blind justice: fairness to groups and the do-no-harm principle, *Journal of Behavioural Decision Making*, Vol. 8, pp. 71-83.
- Bellman, R. (1958), On a routing problem, *Quarterly Applied Mathematics*, Vol. 16, pp. 87-90.
- Bertsekas, D. P. (1991a), An auction algorithm for shortest paths, *SIAM Journal on Optimization*, Vol. 1, pp. 425-447.
- Bertsekas, D. P. (1991b), *Linear network optimization: Algorithms and codes*, MIT Press, Cambridge, MA, U.S.A.
- Bertsekas, D. P. and D. A. Castanon (1991), Parallel synchronous and asynchronous implementations of the auction algorithm, *Parallel Computing*, Vol. 17, pp. 707-732.
- Bertsekas, D. P. and J. N. Tsitsiklis (1989), *Parallel and Distributed Computation*, Prentice Hall, Englewood Cliffs.
- Bertsekas, D. P. and J. N. Tsitsiklis (1991), Some aspects of parallel and distributed iterative algorithms - a survey, *Automatica*, Vol. 27, No. 1, pp. 3-21.

## References

- Bertsekas, D. P., Guerriero, F., and R. Musmanno (1996), Parallel asynchronous label-correcting methods for shortest paths, *Journal of Optimization Theory and Applications*, Vol. 88, Issue 2, pp. 297-320.
- Bitz, F. and H. T. Kung (1988), Path Planning on the WARP Computer: Using a Linear Systolic Array in Dynamic Programming, *International Journal of Computer Mathematics*, Vol. 25, pp. 173-188.
- Bonsall, P. W. and I. Palmer (1997), Do time-based road-user charges induces risk-taking? – Results from a driving simulator, *Traffic Engineering and Control*, Vol. 38, Issue 4, pp. 200-204.
- Bonsall, P., Cho, H.-J., Palmer, I. and N. Thorpe (1998), Experiments to determine drivers' response to road user charges, *Proceedings of European Transport Conference: Seminar D*, PTRC, U.K.
- Bonsall, P. and H.-J. Cho (1999), Travellers' response to uncertainty: the particular case of drivers' response to imprecisely known tolls and charges, *Proceedings of European Transport Conference: Seminar F*, U.K.
- Boot, J., Boot, P. and E. T. Verhoef (1999), *The long road towards the implementation of road pricing: the Dutch experience*, Paper prepared for the ECMT/OECD workshop on managing car use for sustainable urban travel, Dublin, Ireland.
- Brady Shipman Martin, Kirk McClure Morton, Fitzpatrick Associates, and, Colin Buchanan and Partners (1999), *Strategic Planning Guidelines for the Greater Dublin Area*, Department of the Environment and Local Government, Dublin, Ireland.
- Byrne, I. (2002), Electronic Notes, Senior Executive Transportation Planner, Dublin Transportation Office, Dublin 2, Ireland.
- Catling, I. and B. Harbord (1985), Electronic road pricing in Hong Kong: the technology, *Traffic Engineering and Control*, Vol. 26, pp. 608-615.
- Central Statistics Office (2002), *Population statistics from census data*, Ardee Road, Rathmines, Dublin 6, Ireland.
- Cole, S., (1998), *Applied Transport Economics: Policy, Management and Decision Making*, 2<sup>nd</sup> edition, Kogan Page Ltd., London.
- CfIT/MORI (2000), *The CfIT Report: Public Attitudes to Transport in England*, Commission for Integrated Transport, U.K. (available from [www.cfit.gov.uk](http://www.cfit.gov.uk)).



## References

- CfIT/MORI (2001), *The CfIT Report 2001: Public Attitudes to Transport in England*, Commission for Integrated Transport, U.K. (available from [www.cfit.gov.uk](http://www.cfit.gov.uk)).
- Chabini, I. (1998), Discrete dynamic shortest path problems in transportation applications: complexity and algorithms with optimal run time, *Transportation Research Record*, 1645, pp. 170-175.
- Chabini, I., Florian, M. and E. Le Saux (1997), *Parallel and Distributed Computation of Shortest Routes and Network Equilibrium Models*, Dept. of Civil and Environmental Engineering, MIT, Cambridge, MA 02139-4307, U. S. A.
- Chandy, K. M. and J. Misra (1982), Distributed Computation on Graphs: Shortest Path Algorithms, *Comm. ACM*, Vol. 25, pp. 833-837.
- Cherkassky, B. V., Goldberg, A. V. and T. Radzik (1996), Shortest path algorithms: Theory and experimental evaluation, *Mathematical Programming*, Vol. 73, pp. 129-174.
- Chin, K. K. (2002), *Road Pricing: Singapore's Experience*, Paper prepared for the twenty third seminar of the IMPRINT-EUROPE Thematic Network: 'Implementing Reform on Transport Pricing: Constraints and Solutions: learning from best practice', Brussels, 23Rd-24<sup>th</sup> October.
- Clark, D. J., Blythe, P. T., Thorpe, N. and A. Rourke (1994) 'Automatic debiting and electronic payment for transport – the ADEPT project: 3. Congestion metering: the Cambridge trial', *Traffic Engineering and Control*, Vol. 35, pp. 256-63.
- Crauser, A., Mehlhorn, K., Meyer, U., and P. Sanders (1998), *A parallelisation of Dijkstra's shortest path algorithm*, in 23rd Symposium on Mathematical Foundations of Computer Science, LNCS, Brno, Czech Republic, Springer.
- Daniel, J. I. and K. Bekka (2000), The environmental impact of highway congestion pricing, *Journal of Urban Economics*, Vol. 47, pp. 180-215.
- Dawson, J. A. L. and I. Catling (1986), Electronic road pricing in Hong Kong, *Transportation Research Part A: Policy and Practice*, Vol. 20A, No. 2, pp. 129-134.
- Decorla-Souza, P. (1993), Congestion pricing issues and opportunities, *ITE Journal*, Vol. 63, Issue 12, pp. 27-32.
- DEGW (2000), *Managing Intensification and Change: A Strategy for Dublin Building Height*, DEGW plc Porters, North 8 Crinan Street, London N1 9SQ.

## References

- DEHLG (Department of the Environment, Heritage and Local Government) (2000), Local Government (Planning and Development) Act 2000 (available from: <http://www.environ.ie>).
- DEHLG (Department of the Environment, Heritage and Local Government) (2002), National Spatial Strategy (available from: <http://www.environ.ie>).
- DEHLG (Department of the Environment, Heritage and Local Government) (2002), National Sustainable Development Strategy: 1997-2002 (available from: <http://www.environ.ie>).
- DELG & DPE (Department of the Environment and Local Government and the Department of Public Enterprise) (2001), New Institutional Arrangements for Land Use and Transport in the Greater Dublin Area (available from: <http://www.irlgov.ie>).
- Denning, P. J. (1980), Working Sets Past and Present, *IEEE Transactions on Software Engineering*, SE Vol. 6, No. 1, pp. 64-84.
- Deo, N. and C. Pang (1984), Shortest path algorithms: Taxonomy and annotation, *Networks*, Vol. 14, pp. 275-323.
- Deo, N., Pang, C. Y. and R. E. Lord (1980), *Two parallel algorithms for shortest path problems*, Tech Report CS-80-059, Washington State Univ., U.S.A.
- Dey, S. and P. K. Srimani, Fast parallel algorithm for all-pairs shortest path problem and its VLSI implementation, *IEE Proceedings*, Vol. 136, Pt. E, No. 2, March 1989.
- Dijkstra, E. W. (1959), A note on two problems in connection with graphs (spanning tree, shortest path), *Numerical Mathematics*, Vol. 1, pp. 269-271.
- Dix, M. (2002), *The Central London Congestion Charging Scheme – From Conception to Implementation*, Paper prepared for the second Seminar of the IMPRINT-EUROPE Thematic Network: ‘Implementing Reform on Transport Pricing: Identifying Mode-Specific issues’, Brussels, 14<sup>th</sup> / 15<sup>th</sup> May.
- DPE (Department of Public Enterprise) (2000), *A New Institutional and Regulatory Framework for Public Transport*, The Stationary Office, Sun Alliance House, Molesworth Street, Dublin 2, Ireland.
- DTO (Dublin Transportation Office) (2002), *Annual Report*, Dublin Transportation Office, Dublin, Ireland.

## References

- DTO (Dublin Transportation Office) (2001), *A Platform for Change: Final Report*, Dublin Transportation Office, Dublin, Ireland.
- DTO (Dublin Transportation Office) (2000), *A Platform for Change*, Dublin Transportation Office, Dublin, Ireland.
- DTO (Dublin Transportation Office), (1998), *Transportation Review and Short-Term Action Plan*, DTO, Dublin, Ireland.
- Dublin Crisis Conference Committee (1986), *A report on the Dublin Crisis Conference*, Dublin: Published on behalf of the conference organisers by D. Kelly.
- Dublin Transportation Initiative (1995), *Final Report, Technical Annex on Transportation Modelling*, Department of the Environment, Dublin.
- Eager, D. L., Zahorjan, J., and E. D. Lazowska (1989), Speedup versus efficiency in parallel systems, *IEEE Transactions on Computers*, Vol. 38, No. 3, pp. 408-423.
- Edinburgh City Council (2001), *Integrated Transport Initiative for Edinburgh and South-east Scotland*, Appraisal summary table and technical report, October (available from <http://download.edinburgh.gov.uk>).
- Eliasson, J. and L.-G. Mattsson (2001), Transport and location effects of road pricing: A simulation approach, *Journal of Transport Economics and Policies*, Vol. 35, pp. 417-456.
- Emmerink, R. H. M., P. Nijkamp and P. Rietveld (1995), 'Is congestion pricing a first-best strategy in transport policy? A critical review of arguments', *Environment and Planning B: Planning and Design*, Vol. 22, No. 5, pp.581-602.
- ESRI (Economic and Social Research Institute) (2001), *Medium Term Review 2001–2007*, 4 Burlington Road, Dublin 4 (available from: <http://www.ESRI.ie>).
- EUROPRICE (2002), *Technical and Operative Issues: Guidance Paper 3* (available from <http://www.europrice-network.org>).
- EUROPRICE Newsletter (2002), *Reports from cities*, Phase 2, Issue 3 (available from <http://www.europrice-network.org>).
- Ferreira (1981), The role of comprehensive traffic management in energy conservation, PTRC Summer Annual Meeting.

## References

- Fielding, G. J. and D. Klein (1993), *High occupancy/toll lanes: phasing in congestion pricing a lane at a time*, Policy Study No. 170, Los Angeles: Reason Foundation.
- Florian, M. (2001), Applications of parallel computing in transportation: Guest editorial, *Parallel Computing*, Vol. 27, pp. 1521-1522.
- Flowerdew, A. D. J. (1993), *Urban Traffic Congestion in Europe: Road Pricing and Public Transport Demand*, Research Report, The Economist Intelligence Unit, London, May.
- Foo, T. S. (1997), An effective demand management instrument in urban transport: the Area Licensing Scheme in Singapore, *Cities*, Vol. 14, Issue 2, pp. 155-164.
- Foo, T. S. (2000), An advanced demand management instrument in urban transport: electronic road pricing in Singapore, *Cities*, Vol. 17, Issue 1, pp. 33-45.
- Ford, L. R. and D. R. Fulkerson (1962), Maximal flow through a network, *Flows in Networks*, pp. 399-404, Princeton University Press, U.S.A.
- Frank M. and P. Wolfe (1956), An algorithm for quadratic programming, *Naval Research Logistics Quarterly*, Vol. 3, pp. 95-110.
- Gallo, G. and S. Pallottino (1986), Shortest Path Methods: A Unifying Approach, *Mathematical Programming Study*, Vol. 26, pp. 38-64.
- Gallo, G. and S. Pallottino (1984), *Shortest path methods in transportation models*, in (M. Florian, ed.), *Transportation planning models*, North-Holland, pp. 227-256.
- Geist, A., Beguelin, A., Dongarra, J., Jiang, W., Manchek, B. and V. Sunderam, "PVM: Parallel Virtual Machine---A User's Guide and Tutorial for Network Parallel Computing", MIT Press, 1994.
- Gärling, T., A. Gärling and A. Johansson (2000), "Household choices of car-use reduction measures", *Transportation Research Part A: Policy and Practice*, Vol. 34, No. 5, pp. 309-320.
- Gibbons, E (1998), *Application of an Economics Model to Transport Externalities in Dublin*, MSc Dissertation, Department of Civil, Structural and Environmental Engineering, Trinity College, Dublin, Ireland.
- Giuliano, G. (1992), An assessment of the political acceptability of congestion pricing, *Transportation*, Vol. 19, pp. 335-358.

## References

- GLA (2001), *The Mayor's Transport Strategy*, Greater London Authority (GLA), U.K. (available from [www.london.gov.uk/major/strategies/transport/](http://www.london.gov.uk/major/strategies/transport/)).
- Glaister, S. and D. Graham (2000), *The Effect of Fuel Prices of Motorists*, AA Motoring Policy Unit and the United Kingdom Petroleum Industry Association, U.K.
- Glazer, A., Link, H., May, T., Milne, D. and E. Niskanen (2001), *Barriers to transport pricing – review of research*, Paper prepared for IMPRINT-EUROPE Seminar, Brussels 21-22 November, 2001.
- Goh, M. (2002), Congestion management and electronic road pricing in Singapore, *Journal of Transportation Geography*, Vol. 10, pp. 29-38.
- GOL (2000), *Road Charging Options for London: A Technical Assessment*, ROCOL Report, Government Office for London, U.K. (available from [www.go-london.gov.uk](http://www.go-london.gov.uk)).
- Goldberg, A. V. and T. Radzik (1993), A heuristic improvement of the Bellman-Ford algorithm, *Applied Mathematics Letters*, Vol. 6, pp. 3-6.
- Golden, B. (1976), Shortest-path algorithms: A comparison, *Operations Research*, Vol. 24, No. 6, pp. 1164-1168.
- Gomez-Ibáñez, J. A. and G. R. Fauth (1980), Downtown auto restraint policies: the costs and benefits for Boston, *Journal of Transport Economics and Policy*, Vol. 14, pp. 133-53.
- Gómez-Ibáñez, J. A. and K. A. Small (1995), *Road Pricing for Congestion Management: A Survey of International Practice*, Transportation Research Board, National Co-operative Highway Research Program; Synthesis of Highway Practice 210.
- Goodwin, P. (1989), 'The 'rule of three'. A possible solution to the political problem of competing objectives for road pricing', *Traffic Engineering and Control*, October, pp. 495-497.
- Goodwin, P. B. (1992), A review of new demand elasticities with special reference to short and long run effect of price changes, *Journal of Transport Economics and Policy*, Vol. 26, Issue 2, pp. 155-169.
- Goodwin, P. B. (1997), *Solving Congestion*, Inaugural lecture for the professorship of transport policy, University College London, 23<sup>rd</sup> October (available from <http://www.ucl.ac.uk/transport-studies/tsu/pbginau.htm>).

## References

- Goodwin, P. B. and P. M. Jones (1989), *Road pricing: the political and strategic possibilities*, in Systems of Road Infrastructure Cost Coverage, Round Table Report 80, European Conference of Ministers of Transport, Paris, 9-10 February.
- Goodwin, P. B., Hallett, S., Kenny, R. and G. Stokes (1991), *Transport – the New Realism*, Transport Studies Unit, University of Oxford, U.K.
- Greenwood, D. G. and N. B. Taylor (1993), *Parallelising the CONTRAM traffic assignment model*, in European Multi-simulation Conference, Lyon, pp. 7-9.
- Grieco, M. and Jones, P. M. (1994), 'A change in the policy climate? Current European perspectives on road pricing', *Urban Studies*, Vol. 31, Issue 9, pp. 1517-32.
- Gropp, W. and E. Lusk (1997), 'Why Are PVM and MPI So Different?', in Bubak, M., Dongarra, J. (Eds.): *Recent Advances in PVM and MPI*, LNCS, Vol. 1332, Springer-Verlag, Berlin Heidelberg New York, pp. 3-10.
- Gropp W., Lusk, E., and A. Skjellum (1994), *Using MPI – Portable Parallel Programming with the Message-Passing Interface*, The MIT Press, Cambridge, MA, London, England.
- Gropp, W., Huss-Lederman, S., Lumsdaine, A., Lusk, E., Nitzberg, B., Saphir, W., and M. Snir (1998), *MPI: The Complete Reference Volume 2 - The MPI-2 Extension*, MIT Press.
- Gueller, P. (2000), *Surveys, Interviews and Media Analysis – Structure, Questionnaires and Results*, Deliverable number 2 of the PRIMA project (available from [www.certu.fr/internat/peuro/prima/prima.htm](http://www.certu.fr/internat/peuro/prima/prima.htm)).
- Gurd, J., Kirkham, C. C. and I. Watson (1985), The Manchester Prototype Dataflow Computer, *Communications of the Association for Computing Machinery*, Vol. 28, No. 1, pp. 34-52.
- Habbal, M., Koutsopoulos, H. and S. Lerman (1994), A decomposition algorithm for the all-pairs shortest path problem on massively parallel computer architectures, *Transportation Science*.
- Halcrow Fox and Associates (1993), *Road Pricing in London: Review and Specification of Demand Model Elasticities*, a report commissioned by the U.K. Department of Transport, Accent Market Research, University of Leeds, U.K.

## References

- Harrington, W., Krupnick, A. J. and A. Alberini (2001), Overcoming public aversion to congestion pricing, *Transportation Research Part A: Policy and Practice*, Vol. 35, pp. 93-111.
- Harrison, B. (1986), Electronic road pricing in Hong Kong: estimating and evaluating the effects, *Traffic Engineering and Control*, Vol. 27, pp. 13-18.
- Harrison, W. J., Pell, C., Jones, P. M. and H. Ashton (1986), Some advances in model design developed for the practical assessment of road pricing in Hong Kong, *Transportation Research Part A: Policy and Practice*, Vol. 20A, No. 2, pp. 135-143.
- Hårsman, B., Pädam, S. and B. Wijkmark (2000), *Ways and Means to Increase the Acceptance of Urban Road Pricing*, Deliverable 4 of the PRIMA project (available from [www.certu.fr/internat/peuro/prima/prima.htm](http://www.certu.fr/internat/peuro/prima/prima.htm)).
- Hau, T. D. (1990), Electronic road pricing: developments in Hong Kong 1983-1989, *Journal of Transport Economics and Policy*, pp. 203-214.
- Hau, T. D. (1992a), *Economic Fundamentals of Road Pricing: a Diagrammatic Analysis*, World Bank working papers WPS 1070, Report No. TWU1 (available from [http://www.worldbank.org/html/fpd/transport/publicat/pub\\_tran.htm](http://www.worldbank.org/html/fpd/transport/publicat/pub_tran.htm))
- Hau, T. D. (1992b), *Congestion Charging Mechanisms for Roads: an Evaluation of Current Practice*, World Bank working paper WPS-1071, Washington, D.C., Report No. TWU2 (available from [http://www.worldbank.org/html/fpd/transport/publicat/pub\\_tran.htm](http://www.worldbank.org/html/fpd/transport/publicat/pub_tran.htm)).
- Hau, T. D. (1998), 'Congestion pricing and road investment', in *Road Pricing, Traffic Congestion and the Environment* by K. J. Button and E.T. Verhoef, eds., Edward Elgar Publishing Limited, Cheltham, England, pp. 39-78.
- Heaney, Q. R. D. (2001), *Urban development: Use of a Neural Network Mode Choice Model*, PhD Thesis, Department of Civil, Structural and Environmental Engineering, Trinity College, Dublin, Ireland.
- Hempel, R. and D. W. Walker (1999), The emergence of the MPI message passing standard for parallel computing, *Computer Standards and Interfaces*, Vol. 21, pp. 51-62.
- High Level Group on Transport Infrastructure Charging (1999), *Final Report on Estimating Transport Costs* (available from <http://europa.eu.int>).

## References

- Hislop, A. D., McDonald, M. and N. Hounsell (1991), The application of parallel processing to traffic assignment or use with route guidance, *Traffic Engineering and Control*, November, pp. 510-515.
- Hoare, C. A. R. (1978), Communicating Sequential Processes, *Communications of the ACM*, Vol. 21, No. 8, pp. 666-677.
- HOK (Hamilton Osborne King), *Office Surveys*, Hamilton Osborne King, Dublin, Ireland.
- Holland, E. P. and P. L. Watson (1978), Traffic restraint in Singapore, *Traffic Engineering and Control*, Vol. 19, Issue 1.
- Hong Kong SAR Government (2001), *Electronic Road Pricing Feasibility Study: Executive Summary*, Transport Department.
- Hribar, M. R., Taylor, V. E. and D. E. Boyce (1995), *Choosing a shortest path algorithm*, Technical Report CSE-95-004, Computer Science and Engineering, EECS Department, Northern University.
- Hribar, M. R., Taylor, V. E. and D. E. Boyce (2001), Implementing parallel shortest path for parallel transportation applications, *Parallel Computing*, Vol. 27, pp. 1537-1568.
- Hribar, M. R., Taylor, V. E. and D. E. Boyce (1998), Termination detection for parallel shortest path algorithms, *Journal of Parallel and Distributed Computing*, Vol. 55, pp. 153-165.
- Institute for European Environmental Policy (1997), *Road User Exposure to Air Pollution: A Literature Review*, Environmental Transport Association (ETA) on behalf of the Department of Environment, Transport and the Regions, U.K.
- Ison, S. (1996), 'Pricing road space: back to the future? The Cambridge Experience', *Transport Reviews*, Vol. 16, Issue 2, pp. 109-26.
- Ison, S. (1998), A concept in the right place at the wrong time: congestion metering in the city of Cambridge, *Transport Policy*, Vol. 5, pp. 139-146.
- Jaensirisak, S. (2002), *Road User Charging: Acceptability and Effectiveness*, PhD Thesis, University of Leeds, Institute for transport Studies.
- Johansson, B. and L.-G. Mattsson (1995), 'From theory and policy analysis to the implementation of road pricing: the Stockholm region in the 1990s', in Börje Johansson



## References

- and Lars-Göran Mattsson (eds.), *Road Pricing: Theory, Empirical Assessment and Policy*, Boston: Kluwer Academic Publishers, pp. 181-204.
- Jakobsson, C., S. Fuji and T. Gärling (2000), Determinants of private car users' acceptance of road pricing, *Transport Policy*, Vol. 7, pp. 153-158.
- Jones, P. (1991), Gaining public support for road pricing through a package approach, *Traffic Engineering and Control*, Vol. 32, No. 4, pp. 194-196.
- Jones, P. (1998), 'Urban road pricing: public acceptability and barriers to implementation', in K.J. Button and E.T. Verhoef (eds.), *Road Pricing, Traffic Congestion and the Environment*, Edward Elgar, USA.
- Jones, P. and A. Hervik (1992), Restraining car traffic in European cities: an emerging role for road pricing, *Transportation Research Part A: Policy and Practice*, Vol. 26, Issue 2, pp. 133-145.
- Jong, G. de, Biggiero, L., Coppola, P., Gunn, H. F., Dohmen, R., Tegge, O., Palma, A. de, Villemeur, E. de, Duchateau, H. and S. Gayda (1999), *TRACE: Costs of private road travel and their effects on demand, including short and long term elasticities; Deliverable 5: Elasticity Handbook: Elasticities for Prototypical Contexts*, Hague Consulting Group (available from: <http://www.hcg.nl/projects/trace>).
- Killen, J. E. (1992), Transport in Dublin: past, present and future, pp. 305-26, in Halen, F. H. A. and Whelan, K., *Dublin City and Council: From pre-history to present*, Geography Publications, Dublin.
- Knight, F. H. (1924), Some fallacies in the interpretation of social cost, *Quarterly Journal of Economics*, Vol. 38, pp. 582-606.
- Kumar, V. and V. Singh (1991), Scalability of Parallel Algorithms for the All-Pairs Shortest Path Problem, *Journal of Parallel and Distributed Processing*, Vol. 13, No. 2, pp. 124-138.
- Langmyhr, T. (1997), Managing equity: The case of road pricing, *Transport Policy*, Vol. 4, Issue 1, pp. 25-39.
- Larsen, O. I. (1988), The toll ring in Bergen, Norway – the first year of operation, *Traffic Engineering and Control*, Vol. 29, pp. 216-22.
- Larsen, O. I. (1995), The toll cordons in Norway: an overview, *Journal of Transport Geography*, Vol. 3, No. 3, pp. 187-197.

## References

- Lee, G., Kruskal, C. P. and D. J. Kuck (1985), An Empirical Study of Automatic Restructuring of Nonnumerical Programs for Parallel Processors, Special Issue on Parallel Processing of *IEEE Transactions on Computers, Part C*, Vol. 34, No. 10, pp. 927- 933.
- Leonard, D. R. and P. Gower (1989), *User Guide to CONTRAM Version 4*, TRRL Supplementary Report 735, Transport and Road Research Laboratory, Crowthorne.
- Leonard, D. R., Gower, P., and N. B. Taylor (1989), *CONTRAM: Structure of the Model*, TRL Report RR178, Transport Research Laboratory, Crowthorne, U.K.
- Lewis, N. C. (1993), *Road Pricing: Theory and Practice*, London: Thomas Telford, U.K.
- LEX Service (1998), *The 1998 Lex Report on Motoring: Driving for the Future*, U.K.
- LEX Service (1999), *The 1999 Lex Report on Motoring: the Authoritative Research on the Attitudes and Behaviour of Britain's Motorists*, U.K.
- Litman, T. (1999), *Using Road Pricing Revenue: Economic Efficiency and Equity Considerations*, Victoria Transport Policy Institute, Canada.
- London First (1999), *Transport Initiative – Road User Charging – A Proposal for Central London*, London First, U.K.
- London Planning Advisory Committee (1988), *Strategic Planning Advice for London*, London (Cited in: Dix, 2002).
- Luk, J. (1999), Electronic road pricing in Singapore, *Road & Transport Research*, Vol. 8, Issue 4, pp. 28-40.
- Luk, J. and S. Hepburn (1993), *New Review of Australian Travel Demand Elasticities*, Research Report (ARR 249), ARRB Transport Research Ltd, Australia.
- Luk, J. and E. Chung (1997), *Public Acceptance and Technologies for Road Pricing*, Research Report (ARR 307), ARRB Transport Research Ltd, Australia.
- MacLaran, A. (1993), *Dublin: the shaping of a capital*, London: Belhaven Press.
- Mateti, P. and N. Deo (1982), Parallel Algorithms for the Single Source Shortest Path Problem, *Computing*, Vol. 29, pp. 31-49.

## References

- Mattsson, L.-G. (2003a), *Modelling road pricing reform in Stockholm*, Royal Institute of Technology, Department of Infrastructure, Division of Systems Analysis and Economics, Unit for Transport and Location Analysis, SE-100 4 Stockholm, Sweden.
- Mattsson, L.-G. (2003b), *Road pricing: Consequences on traffic, congestion and location*, Paper prepared for the research conference: 'The Economic and Environmental Consequences of Regulating Traffic', Hillerod, Denmark, February 2-3.
- Mattsson, L.-G. and L. Sjölin (2002), *Transport and location effects of a ring road with or without road pricing*, Royal Institute of Technology, Department of Infrastructure, Division of Systems Analysis and Economics, Unit for Transport and Location Analysis, SE-100 4 Stockholm, Sweden.
- Matzoros, T., Van Vliet, D., Randle, J. and B Weston (1987), A validation of the SATURN and ME2 models using before-and-after survey data from Manchester, *Traffic Engineering and Control*, Vol. 28, pp. 641-643.
- May, A. D. (1975), Supplement licensing: an evaluation, *Traffic Engineering and Control*, Vol. 16, No. 4, pp. 162-7.
- May, A. D. (1992), Road pricing: an international perspective, *Transportation*, Vol. 19, Issue 4, pp. 313-333.
- May, A. D., Guest, P.W. and K. Gardner (1990), Can rail-based policies relieve urban traffic congestion?, *Traffic Engineering and Control*, Vol. 31, pp. 406-7.
- May, A D, Milne, D S, van Vliet, D, Smith, M J and M Wisten (1994), *The network effects of alternative road pricing systems*, in Proceedings of the 7th IEE Conference on Road Traffic Monitoring and Control (Pub No 391), pp. 46-49.
- May, A. D., P. W. Bonsall and P. J. Hills (1998), *The impacts of different road user charging systems*, in Proceedings of the International Conference on Transportation into the Next Millennium, Nanyang Technological University, Singapore, 9-11.
- May, A. D. and D. S. Milne (2000), Effects of alternative road pricing systems on network performance, *Transportation Research Part A: Policy and Practice*, Vol. 34, Issue 6, pp. 407-430.
- May, A. D., Shepherd, S. P. and J. J. Bates (2000), Supply curves for urban road networks, *Journal of Transport Economics and Policy*, Vol. 34, Issue 3, pp. 261-290.
- McCarthy, P. and R. Tay (1993), Pricing road congestion: Recent evidence from Singapore, *Policy Studies Journal*, Vol. 21, Issue 2, pp. 296-308.

## References

- Meland, S. and J. Polak (1993), *Impacts of the Trondheim toll ring on travel behaviour: some preliminary findings*, Proceedings 21<sup>st</sup> PTRC Summer Annual Meeting: Seminar F, pp. 103-115.
- Menon, A. P. G. (2000), EPR in Singapore-a perspective one year on, *Traffic Engineering and Control*, Vol. 41, No. 2, pp. 40-44.
- Menon, A. P. G. and K. K. Chin (1998), *The making of Singapore's electronic road pricing system*, in Proceedings of the International Conference on Transportation into the Next Millennium, Centre for Transportation Studies, Nanyang Technological University, Singapore (September 9-11)
- Menon, A. P. G. and S. H. Lam (1993), 'Singapore's road pricing systems, 1989-1993', Transportation Research Report NTU/CTS/93-2, Centre for Transportation Studies, Nanyang Technological University, Singapore (November) (Cited in: Small and Gomez-Ibanez, 1998).
- Menon, A. P. G., Lam, S. H. and H. S. L. Fan (1993), 'Singapore's road pricing scheme: Its past, present and future', *ITE Journal*, pp. 44-48 (Cited in: Phang and Toh, 1997).
- Message Passing Interface Forum (1994), MPI: A Message Passing Interface Forum, *International Journal of Supercomputer Applications and High Performance Computing*, Vol. 8, No. 3/4, pp. 165-414.
- Minister Martin Cullen TD (2002), *Government Launches National Spatial Strategy - 28 November*, Department of the Environment, Heritage and Local Government, Media Centre, Press Releases (available from: <http://www.environ.ie>).
- Minsky, M. and S. Papert (1971), On some associative, parallel and analog computations, in *Associative Information Technologies*, E. J. Jacks, Ed., New York: Elsevier North Holland.
- Mohring, H. (1963), Urban highway investments, in Rover Dorfman (ed.), *Measuring Benefits of Government Investment*, Washington, DC: The Brookings Institution, pp. 231-75.
- Mondou, J. F., Crainic, T. G. and S. Nguyen (1991), Shortest Path Algorithms: A Computational Study with C Programming Language, *Computers and Operations Research*, Vol. 18, pp. 767-786.

## References

- Moore, E. F. (1957), *The shortest path through a maze*, in Proceedings of an international symposium on the theory of switching, pp. 285-292, Harvard University.
- Morgan, R. T. and Associates (1973), *The Central Dublin Traffic Plan*, Dublin: Corporation of Dublin, Ireland.
- Mu-consult (1998), *Varianilisatie van de autokosten*, Amersfoort, (Cited in: Boot et al., 1999)
- MVA (1995), *The London Congestion Charging Research Programme: Principal Findings*, The MVA Consultancy, HMSO: London.
- Nagel, K. and M. Rickert (2001), Parallel implementation of the TRANSIMS micro-simulation, *Parallel Computing*, Vol. 27, No. 12, pp. 1611-1639.
- Narayanan, P. (1992), *Single source shortest path problem on processor arrays*, in Proceedings Frontiers of Massively Parallel Computation, pp. 553-556, McLean, VA.
- NEDO (1991), *A Road User Charge? Londoner' View*, National Economic Development Office, U.K., (Cited in: Flowerdew, 1993).
- NERA (2002), *Paying for Road Use: Technical Report*, a report to the Commission for Integrated Transport (CfIT), National Economic Research Associates (NERA), U.K. (available from [www.cfit.gov.uk](http://www.cfit.gov.uk)).
- New Nation (1975), *New Time Limits*, July 10 (Cited in: Phang and Toh, 1997)
- Newbery, D. M. (1990), Pricing and congestion: economic principles relevant to pricing roads, Special issue on Transport, *Oxford Review of Economic Policy*, Vol. 6, 22-38
- Newbery, D. M. (2002), *Road user and congestion charges*, Paper prepared for Excise Tax Policy and Administration Conference, OCFEB (Research Centre for Economic Policy, Erasmus University) Ministry of Finance, The Hague, 11-12 April.
- Nilsson, M. and R. Kuller (2000), Travel behaviour and environmental concern, *Transportation Research Part D*, Vol. 5, pp. 211-234.
- O'Mahony, M., Geraghty, D. and I. Humphreys (2000), Distance and time based road pricing trial in Dublin, *Transportation*, Vol. 27, pp. 269-283.
- Odeck, J. and S. Brathen (1997), On public attitudes toward implementation of toll roads – the case of Oslo toll ring, *Transport Policy*, Vol. 4, Issue 2, pp. 73-83.

## References

- Odeck, J. and S. Brathen (2002), Toll financing in Norway: The success, the failures and perspectives for the future, *Transport Policy*, Vol. 9, pp. 253-260.
- Oldbridge, B. (1995), Congestion metering in Cambridge City, United Kingdom, in B. Johansson and L. Mattson (eds.), *Road Pricing: Theory, Empirical Assessment and Policy*, Kluwer Academic Publishers, USA.
- Oldfield, R. H. and Tyler (1981), *The elasticity of medium distance rail travel, Report LR 993*, Transport Research Laboratory, Crowthorne, Berkshire, U.K.
- Orski, C. (1998), 'Value pricing: paying a premium for congestion-free travel', *Traffic Technology International*, pp. 37-38.
- Ortuzar, J. de D. and L. G. Willumsen (1994), *Modelling Transport*, John Wiley and Sons Ltd., England.
- Oscar Faber Consultants in association with Goodbody Economic Consultants Ltd., TSRG and TORG (1998), *A Study of Road Pricing in Dublin*, Department of the Environment and Local Government, Dublin 1, Ireland.
- Oum, T. H., Waters II, W. G and J. Yong (1992), Concepts of price elasticities of transport demand and recent empirical estimates, *Journal of Transport Economics and Policy*, Vol. 26, pp. 139-154.
- Paige, R. C. and C. P. Kruskal (1985), *Parallel algorithms for shortest path problems*, in IEEE, pp. 14-20.
- Pallottino, S. (1984), Shortest path methods: Complexity, interrelations and new propositions, *Networks*, Vol. 14, pp. 257-267.
- Pallottino, S. and M. Scutella (1997), *Shortest path algorithms in transportation models: Classical and innovative aspects*, Technical Report: TR-97-06, Universit di Pisa.
- Papaefthymiou, M. and J. Rodrigue (1994), *Implementing parallel shortest-paths algorithms*, Department of Computer Science, Yale University.
- Pape, U. (1974), Implementation and efficiency of Moore algorithms for the shortest route problem, *Mathematical Programming*, Vol. 7, pp. 212-222.
- Phang, S. Y. and R. S. Toh (1997), From manual to electronic road congestion pricing: the Singapore experiment, *Transportation Research Part E: Logistics and Transportation Review*, Vol. 33, No. 2, pp. 97-106.

## References

- Pigou, A. C., (1920), *Wealth and Welfare*, MacMillan, U.K.
- Pollock, M. and W. Wiebenson (1960), Solutions for the shortest route problem, *Operations Research*, Vol. 8, pp. 224-230.
- Polymenakos, L., and D. P. Bertsekas (1994), Parallel Shortest Path Auction Algorithms, *Parallel Computing*, Vol. 20, pp. 1221-1247.
- PRIMA (PRicing Measures Acceptance) *Ways and Means to Increase the Acceptance of Urban Road Pricing* (available at: <http://europa.eu.int/comm/transport/extra/primia.html>).
- QBC Steering Committee (2002), QBC Report, prepared for DTO Steering Committee as a response to a request at its meeting of July 2002 for a comprehensive statement on the QBC strategy.
- Ramjerdi, F. and O., Larsen (1991), *Road Pricing as a Means of Financing Investment in Transportation Infrastructure, the Case of Oslo*, in Proceedings of PTRC 19<sup>th</sup> Summer Annual Meeting.
- Raux, C., Andan O. and C. Godinot (1998), 'The simulation of behaviour in an inexperienced future: the case of urban road pricing', in J. Ortuzar, D. Hensher and S. Jara-Diaz (eds.), *Travel Behaviour Research: Updating the State of Play*, Chapter 5, pp. 67-86, Pergamon, U.K.
- Richards, M. and B. Harrison (1999), *Road Pricing in Leeds*, MVA Consultants, U.K.
- Rietveld, P. and E. T. Verhoef (1998), Social feasibility of policies to reduce externalities in transport in K. J. Button and E. T. Verhoef (eds.), *Road Pricing, Traffic Congestion and the Environment*, Edward Elgar, USA.
- Salford Software Ltd. (1996), Salford FTN77 User's Guide, Salford Software: The Compiler Specialists, Adelphi House, Adelphi Street, SALFORD, M3 6EN, U.K.
- Santos, G. (2000a), *On the economic, technological and political aspects of road pricing as a tool for traffic demand management*, Department of Applied Economics, University of Cambridge, CB3 9DE, U.K.
- Santos, G. (2000b), *Possible road pricing and use of revenues in England*, Department of Applied Economics, University of Cambridge, CB3 9DE, U.K.
- Santos, G., Rojey, L. and D. Newbery (2000), *The environmental benefits from road pricing*, Department of Applied Economics, University of Cambridge, U.K.

## References

- Schade, J. (2002), *Public and Political Acceptability as Criteria for Implementation Paths: Urban and Interurban Road*, Paper prepared for Conference on Acceptability of Transport Pricing Strategies, Dresden, Germany, pp. 23-24.
- Schade, J. (2003), European research results on transport pricing acceptability, in J. Schade and B. Schlag (Eds.), *Acceptability of transport pricing strategies*, pp. 109-124, Oxford: Elsevier.
- Schade, J. and B. Schlag (2003), *Acceptability of Pricing Reform*, Paper prepared for the IMPRINT-EUROPE Seminar, 13-14 May.
- Schade, J. and B. Schlag (2000), *Acceptability of Urban Transport Pricing: AFFORD*, VATT Research Projects 72, Helsinki. Funded by the European Commission 4<sup>th</sup> Framework – Transport RTD.
- Schlag, B. and U. Teubel (1997), Public acceptability of transport pricing, *IATSS Research*, Vol. 21, Issue 2, pp. 134-142 (Cited in Schade, 2002).
- Schlag, B. and J. Schade (2000), Public acceptability of traffic demand management and pricing measures in Europe, *Traffic Engineering and Control*, Vol. 41, Issue 8, pp. 314-318.
- Sharpe, J. M. (1993), Demand management: The Cambridge approach *Transportation Studies*, Cambridgeshire County Council, Cambridge, U.K. (Cited in: Small and Gomez-Ibanez, 1998).
- Sheldon, R., Scott, M. and P. Jones (1993), *London congestion charging: exploratory social research among London residents*, Paper presented at the 21<sup>st</sup> PTRC Summer Annual Meeting. In: Proceedings of Seminar F, PTRC Education and Research Services, Glenthorne House, Hammersmith Grove, London W4, pp. 129-145.
- Shiftan, Y., Shefer, D. and G. Avraham (2003), *The impact of auto restrain policies on the location of businesses in city centres*, Transportation research Board Annual Meeting CD-ROM.
- Small, K. A. (1992), Using the revenues from congestion pricing, *Transportation*, Vol. 19, pp. 359-381.
- Small, K. A., Supernak, J., Collins, T. and E. Schreffler (1997), *I-15 Congestion Pricing Project, Monitoring and Evaluation Services: Task 1.2*, Up-to-date results of the SR-91 congestion pricing experiment, prepared for San Diego Association of Governments.



## References

- Small, K. A. and J. A. Gomez-Ibanez (1998), Road pricing for congestion management: the transition from theory to policy in K. J. Button and E. T. Verhoef (eds.), *Road Pricing, Traffic Congestion and the Environment*, Edward Elgar Publishers, USA, pp. 213-246.
- Smith, J.C. and R. S. Burton (1998), *Initial findings from the Leicester environmental road tolling scheme*, paper presented at the 9<sup>th</sup> International Conference on Road Transport Information & Control, IEE, pp. 21-23, London, U.K.
- Snir, M., Otto, S. W., Huss-Lederman, S., Walker, D. W., and J. Dongarra (1996), *MPI: The Complete Reference*, Scientific and Engineering Computation Series, MIT Press.
- Snir, M., Otto, S., Huss-Lederman, S., Walker, D., and J. Dongarra (1998), *MPI: The Complete Reference Volume 1 - The MPI Core*, 2<sup>nd</sup> Ed., MIT Press.
- Sociaal-Economische Raad (The Social and Economic Council) (2001), *National Traffic and Transport Plan 2001 – 2020*, Bezuidenhoutseweg 60, PO Box 90405, 2509 LK The Hague, The Netherlands.
- Social Democratic Party, Moderate Party and Liberal Party (1991), *The Greater-Stockholm negotiation on traffic and environment: the Dennis agreement*, signed by O. Lindkvist *et al.*, transmitted by Bengt Dennis to The Minister of Transportation and Communication, Stockholm, Sweden (23 January).
- SPGO (Strategic Planning Guidelines Office) (2002), *Strategic Planning Guidelines for the Greater Dublin Area: Review and Update*, prepared for the Department of the Environment and Local Government and local councils, The Stationary Office, Sun Alliance House, Molesworth Street, Dublin 2, Ireland.
- Steer Davis Gleave in association with McHugh Consultants (1992), *Dublin Transportation Study: Phase 1*, for Dublin Transportation Review Group, Government Stationary Office, Dublin, Ireland.
- Steer Davis Gleave in association with McHugh Consultants (1994a), *Dublin Transportation Initiative (DTI): Final Report*, Government Stationary Office, Dublin, Ireland.
- Steer Davis Gleave in association with McHugh Consultants (1994b), *Dublin Transportation Initiative (DTI): Technical Volume 1 and 2*, Government Stationary Office, Dublin, Ireland.
- Stokes, G. (1996), *Getting car commuters onto public transport – a Survey in Merseyside*, PTRC European Transport Forum (24th: 1996: Brunel University, Seminar

## References

- F). Public transport planning and operations: proceedings of Seminar F, PTRC Education and Research, London, U.K.
- Stokes, G. and B. Taylor (1995), *The public acceptability of sustainable transport policies: finding from the British Social Attitudes survey*, PTRC European Transport Forum (23rd: 1995: University of Warwick, Seminar C). Transport policy and its implementation: proceedings of Seminar C, pp. 69-83, PTRC Education and Research Section, London, U.K.
- Stoneman, A. C. (1998), *Forecasting the impact of demand management strategies: Results from SATURN model tests (part 1)*, Institute for Transport Studies, Leeds, U.K.
- Sullivan, E. C. (1996), *Technical Report No. 3: Evaluating the impacts of the SR 91 variable-toll express lane facility*, presentation of interim findings, Department of Civil and Environmental Engineering, Cal. Poly. State University, San Luis Obispo, (Cited in: Small et al., 1997).
- Sunderam, V., PVM: A Framework for Parallel Distributed Computing, *Concurrency: Practice and Experience*, Vol. 2, No. 4, 315-339, December 1990.
- Taylor, N. B. (1990), *CONTRAM 5: An enhanced traffic assignment model*, TRL Report RR249, Transport Research Laboratory, Crowthorne, U.K.
- Taylor, N.B. (1992), *Dynamic route guidance simulation with CONTRAM and ROGUS*, in Proceedings of PROMETHEUS Workshop on Traffic Related Simulation, Stuttgart.
- Taylor, B. and L. Brook (1998), Public attitudes to transport issues: finding from the british social attitudes surveys, in D. Banister (ed.), *Transport Policy and the Environment*, E&FN SPON, U.K.
- TCRP (Transit Co-operative Research Program) Report 95 (2003), *Traveler Response to Transportation System Changes: Chapter 14 – Road Value Pricing*, Transportation Research Board, Washington, D. C.
- Thompson, A. (1990), Road user charging: the current state of technology, *Traffic Engineering and Control*, Vol. 31, No. 10.
- Thorpe, N., Hills, P. and S. Jaensirisak (2000), Public attitudes to TDM measures: a comparative study, *Transport Policy*, Vol. 7, Issue 4, pp. 243-257.
- Traff, J. L. (1995), An experimental comparison of two distributed single-source shortest path algorithms, *Parallel Computing*, Vol. 21, pp. 1505-1532.

## References

- Transport Bureau (1998), *Feasibility Study on Electronic Road Pricing – Progress Report*, Hong Kong SAR Government.
- Transport Consultative Commission (1980), *Passenger Transport Services in the Dublin Area: Report of the Transport Consultative Commission to the Minister for Transport*, The Stationary Office, Sun Alliance House, Molesworth Street, Dublin 2, Ireland.
- Transport for London (2003), *National Transport Model – Working Paper 3*, Department for Transport, U.K. (available from: <http://www.dft.gov.uk>)
- Transport for London (2001), *Greater London Authority Act 1999*, Department for Transport, U.K. (available from: <http://www.dft.gov.uk>)
- Transportation Research Board: Committee for Study on Urban Transportation Congestion Pricing (1994), *Curbing Gridlock: Peak-Period Fees to Relieve Traffic Congestion*, Special Report 242, Volume 1: Committee Report and Recommendations; Volume 2: Commissioned Papers, National Academy Press, USA.
- Transpotech, Ltd (1985), *Electronic Road Pricing Pilot Scheme: Main Report*, Report prepared for the Hong Kong Government (May).
- TransPrice (1999), *Trans Modal Integrated Urban Transport Pricing for Optimum Modal Split: Final Report for Publication*, Project funded by the European Commission under the Transport RTD Programme of the 4<sup>th</sup> Framework Programme.
- Tremblay, N. and M. Florian (2001), Temporal shortest paths: Parallel computing implementations, *Parallel Computing*, Vol. 27, pp. 1569-1609.
- Tretvik, T. (1992), *The Trondheim toll ring: applied technology and public opinion*, SINTEF Transport Engineering, Trondheim, Norway, Presented at the joint OECD/ECMT/GVF/NFP Conference on the Use of Economic Instruments in Urban Travel Management, Basel, Switzerland (June), (Cited in: Small and Gomez-Ibanez, 1998).
- Tseng, P., Bertsekas, D. P., and J. N. Tsitsiklis (1990), Partially asynchronous, parallel algorithms for network flow and other problems, *SIAM Journal of Control and Optimization*, Vol. 28, pp. 678-710.
- U.K. Ministry of Transport (1964), *Road Pricing: The Economic and Technical Possibilities*, Report of the committee chaired by Smeed, Her Majesty's Stationery Office (HMSO), London.

## References

- U.K. DETR (1998a), *Breaking the Logjam: consultation paper*, The U.K. Government's consultation paper on fighting traffic congestion and pollution, Department of Environment, Transport and the Regions, U.K. (available at <http://www.dft.gov.uk>).
- U.K. DETR (1998b), *Transport and the economy: full report*, Standing Advisory Committee on Trunk Road Assessment (SACTRA), Department of the Environment, Transport and the Regions, U.K. (available at <http://www.dft.gov.uk>).
- U.K. DETR (2001), *Analysis of Congested Network*, project carried out by ITS (University of Leeds), TORG (University of Newcastle upon Tyne), John Bates' Services (Consultant) and TRL (Transport Research Laboratory) (available at <http://www.dft.gov.uk>).
- U.K. DETR (2000), *Improved Elasticities and Methods. Phase 4 Report: Re-analysis of Congestion Data*, carried out by MVA Consultancy and John Bates Services for the Department of the Environment, Transport and the Regions, U.K. (available at <http://www.dft.gov.uk>).
- U.K. DETR (2001), *Development of Integrated Assignment and Demand Modelling*, carried out by WS Atkins Consultants Limited for the Department of the Environment, Transport and the Regions, U.K. (available at <http://www.dft.gov.uk>).
- US EPA (1997), *Opportunities to Improve Air Quality through Transportation Pricing Programs*, U.S. Environmental Protection Agency, Washington, D.C., U.S.A. (available at <http://www.epa.gov/oms/market/pricing.pdf>).
- US EPA (1998), *Technical Methods for Analysing Pricing Measures to Reduce Transportation Emissions*, Washington, D.C., U.S.A.
- Van der Loop, J.T.A. and I.H. Veling (1994), 'Developments in the acceptance of the transport structure plan by the Dutch population and how to increase the acceptance of measures?', Proceedings 22<sup>nd</sup> Summer Annual Meeting: Seminar D, pp. 125-134.
- Van Vliet, D. (1977), D'Esopo: a forgotten tree-building algorithm, *Traffic Engineering and Control*, pp. 372-373.
- Van Vliet D. (1978), Improved shortest path algorithms for transport networks, *Transportation Research*, Vol. 12, pp. 7-20.
- Van Vliet, D. (1982), SATURN – A modern assignment model, *Traffic Engineering and Control*, Vol. 23, No. 12, pp. 578-581.

## References

- Van Vliet, D. and M. Hall (1998), *SATURN 9.4 User Manual*, W.S. Atkins, U.K.
- Verhoef, E. T. (1996), *The Economics of Regulating Road Transport*, Edward Elgar, Cheltenham, U.K.
- Verhoef, E. T. (2003), *Marginal Cost Based Pricing in Transport: Key Implementation Issues from the Economic Perspective*, Paper prepared for the IMPRINT-EUROPE Seminar, 13-14 May.
- Vickrey, W. S. (1955), Pricing in Transportation and Public Utilities, Some implications of marginal cost pricing for public utilities, *The American Economic Review*, Vol. 45, No. 2, Papers and Proceedings of the Sixty-seventh Annual Meeting of the American Economic Association, pp. 605-620.
- Vickrey, W. S. (1963), Pricing and resource allocation in transportation and public utilities, Pricing in Urban and Suburban Transport, *The American Economic Review*, Vol. 53, No. 2, Papers and Proceedings of the Seventy-fifth Annual Meeting of the American Economic Association, pp. 452-465.
- Vickrey, W. S. (1969), Congestion theory and transport investment, *The American Economic Review*, Vol. 59, No. 2, Papers and Proceedings of the Eighty-first Annual Meeting of the American Economic Association, pp. 251-260.
- Vorhees and Associates (1975), *The Dublin Rail Rapid Transit Study*, prepared for Coras Iompair Eireann, Dublin: C.I.E.
- Walker, D. W. (1994), The design of a standard message-passing interface for distributed memory concurrent computers, *Parallel Computing*, Vol. 20, No. 4, pp. 657-673.
- Walters, A. (1954), Track costs and motor taxation, *Journal of Industrial Economics*, Vol. 2, Issue 2, pp. 135-146.
- Walters, A. (1961), The theory and measurement of private and social cost of highway congestion, *Econometrica*, Vol. 29, Issue 4, pp. 676-699.
- Wardrop, J. G. (1952), *Some theoretical aspects of road traffic research*, in Proceedings of the Institution of Civil Engineers, (Part II), Vol. 1, pp. 325-378.
- Watson, P. L. and E. P. Holland (1978), *Relieving Traffic Congestion: The Singapore Area Licensing Scheme*, World Bank Staff Working Paper No. 281, Washington, D.C., U.S.A.

## References

- Willoughby, C. (2001), Singapore's motorization policies 1969-2000", *Transport Policy*, Vol. 8, Issue 2, pp. 125-139.
- Wilson, P. W. (1988), Welfare effects of congestion pricing in Singapore, *Transportation*, Vol. 15, Issue3, pp. 191-210.
- Wright, M. (1967), *The Dublin Region: Advisory Regional Plan and Final Report*, Department of Local Government, Ireland.
- Zenios, S. A. (1989), Parallel numerical optimisation: current status and annotated bibliography, *ORSA Journal on Computing*, Vol. 1, No. 1, pp. 20-43.
- Zettel R. and R. Carll (1964), *The basic theory of efficiency tolls: the tolled, the tolled-off and the untolled*, in Highway Research Record 47, HRB, National Research Council, US Transport Research Board, Washington, D.C., U.S.A.
- Zhan, F. B. and C. E. Noon (1998), Shortest Path Algorithms: An Evaluation using Real Road Networks, *Transportation Science*, Vol. 32, No. 1, pp. 65-73.
- Ziliaskopoulos, A. and D. Kotzinos (2001), A massively parallel time-dependent least-time-path algorithm for intelligent transportation systems applications, *Computer-Aided Civil and Infrastructure Engineering*, Vol. 16, Issue 5, Special Issue: High Performance Computing, pp. 337-346.
- Ziliaskopoulos, A. K. and H. S. Mahmassani (1996), A Note on Least Time Path Computation Considering Delays and Prohibitions for Intersection Movements, *Transportation Research*, Vol. 30, pp. 359-367.
- Ziliaskopoulos, A. and H. S. Mahmassani (1992), *Design and implementation of a shortest path algorithm with time-dependent arc costs*, in Proceedings of the 5<sup>th</sup> Advanced Technology Conference, Washington, D.C., U.S.A., pp. 1072-1093.
- Ziliaskopoulos, A. and H. S. Mahmassani (1993), Time-Dependent, Shortest-Path Algorithm for Real-Time Intelligent Vehicle Highway System Applications, *Transportation Research Record*, Vol. 1408, pp. 94-100.
- Ziliaskopoulos, A., Kotzinos, D. and H. S. Mahmassani (1997), Design and implementation of parallel time-dependent least time path algorithms for intelligent transportation systems applications, *Transportation Research C*, Vol. 5, No. 2, pp. 95-107.