

**DELINEATION OF THE GEOSPATIAL DIMENSIONS OF THE
RESIDENTIAL REAL ESTATE SUBMARKET STRUCTURE**

Anthony J. M. Lockwood
B.App.Sc.(Valuation), Grad.Dip.(App.Stats), MSIS

Thesis submitted for the degree of

Doctor of Philosophy

Department of Geographical and Environmental Studies

School of Social Sciences

The University of Adelaide

December 2007

TABLE OF CONTENTS

TABLE OF CONTENTS	II
LIST OF TABLES	V
LIST OF FIGURES	VI
ABSTRACT	VII
DECLARATION	IX
ABBREVIATIONS AND ACRONYMS	X
ACKNOWLEDGEMENTS	XI
CHAPTER 1 INTRODUCTION	1
1.1 DEFINING THE RESEARCH QUESTION.....	1
1.2 OBJECTIVE OF THE STUDY	4
1.3 NATURE OF SUBMARKETS.....	5
1.3.1 <i>The importance of understanding submarket structure</i>	8
1.3.2 <i>Identified Issues in defining and delineating submarkets</i>	11
1.4 CHOICE OF SUBMARKET DEFINITION.....	11
1.5 ACHIEVING THE OBJECTIVE	13
1.5.1 <i>Stage 1 – The creation of the residential living structure (RLS)</i>	14
1.5.2 <i>Stage 2 – The relationship of the marketplace to the RLS</i>	15
1.6 CONCLUSIONS	18
1.6.1 <i>Contribution of the thesis</i>	19
1.7 STRUCTURE OF THE THESIS.....	20
CHAPTER 2 THE NATURE OF THE RESIDENTIAL REAL ESTATE MARKET AND ITS SUBMARKETS	21
2.1 INTRODUCTION.....	21
2.2 THE GENERAL REAL ESTATE MARKET	22
2.3 THE IMPORTANCE OF IDENTIFYING SUBMARKET BOUNDARIES	27
2.3.1 <i>Housing policy</i>	28
2.3.2 <i>Property taxation</i>	31
2.4 DEFINITION OF A SUBMARKET	32
2.4.1 <i>General comment</i>	32
2.4.2 <i>Substitutability</i>	33
2.4.3 <i>Spatial or aspatial submarket identifiers</i>	35
2.4.4 <i>Specific submarket identifiers</i>	37
2.4.5 <i>Diagrammatic submarket concepts</i>	39
2.4.6 <i>Summary of spatial submarket definitions</i>	42
2.5 ISSUES IDENTIFIED IN THE LITERATURE	46
2.5.1 <i>Opportunity for contribution</i>	46
2.6 SUBMARKET DEFINITION ADOPTED FOR THIS STUDY.....	47
2.6.1 <i>The use of the term ‘submarket’</i>	51
2.7 CONCLUSION	55
CHAPTER 3 THE NATURE OF THE STUDY AREA	57
3.1 INTRODUCTION.....	57
3.2 GEOGRAPHICAL DIMENSIONS	60
3.3 THE RESIDENTIAL STRUCTURE.....	61
3.4 THE RESIDENTIAL MARKET.....	66
3.5 CREATING THE DATA SET	70

3.5.1	<i>The land information management environment</i>	70
3.5.2	<i>Data identification and selection</i>	72
3.5.3	<i>Stage 1</i>	73
3.5.4	<i>Stage 2</i>	79
3.6	DATA SOURCES.....	80
3.6.1	<i>Australian Bureau of Statistics</i>	80
3.6.2	<i>The University of Adelaide</i>	80
3.6.3	<i>The State Government of South Australia</i>	80
3.7	PREPARATION AND DERIVATION OF CERTAIN DATA ATTRIBUTES.....	80
3.7.1	<i>Amenity attribute</i>	81
3.7.2	<i>Socio economic attributes</i>	82
3.7.3	<i>Added value structural attribute</i>	82
3.8	DATA ADOPTED FOR THE STUDY	83
3.8.1	<i>Previous data selection processes</i>	85
3.9	CONCLUSION	86
CHAPTER 4 METHODOLOGY OVERVIEW		88
4.1	INTRODUCTION	88
4.2	ANALYTICAL TECHNIQUES	90
4.2.1	<i>A context</i>	90
4.3	THE SUBMARKET LITERATURE.....	92
4.4	CHOICE OF ANALYTICAL TECHNIQUES	100
4.4.1	<i>Stage 1: Principal component analysis</i>	102
4.4.2	<i>Stage 2: Geographically weighted regression and cluster analysis</i>	104
4.5	COMPUTER SOFTWARE	107
4.6	VERIFICATION OF THE RESULTS.....	108
4.7	CONCLUSION	109
CHAPTER 5 METHODOLOGY: STAGE 1 DERIVATION OF THE RESIDENTIAL LIVING STRUCTURE (RLS)		112
5.1	INTRODUCTION.....	112
5.2	ANALYTICAL STEPS	114
5.2.1	<i>Limitations</i>	120
5.3	CONCLUSION	122
CHAPTER 6 METHODOLOGY: STAGE 2 DERIVATION OF GEOSPATIAL RESIDENTIAL REAL ESTATE SUBMARKET BOUNDARIES		124
6.1	INTRODUCTION.....	124
6.2	DATA SETS	125
6.2.1	<i>Property transaction price</i>	125
6.2.2	<i>Regular 300 metre grid</i>	127
6.3	ANALYSIS TECHNIQUES	127
6.3.1	<i>Geographically weighted regression (GWR)</i>	127
6.3.2	<i>Cluster analysis (CA)</i>	130
6.3.3	<i>Sequence of analytical steps</i>	133
6.4	UNDERSTANDING THE STRUCTURE OF GEOSPATIAL SUBMARKETS.....	140
6.5	CONCLUSION	142
CHAPTER 7 RESULTS OF STAGE 1: DERIVATION OF THE RLS		144
7.1	INTRODUCTION.....	144
7.2	SELECTION OF THE NUMBER OF FACTORS	145
7.2.1	<i>The scree plot</i>	146
7.2.2	<i>Sensibility of factors</i>	146
7.2.3	<i>Communality</i>	147
7.3	THE CHOICE OF ROTATION – ORTHOGONAL OR OBLIQUE	148
7.4	THE INTERPRETATION OF THE FACTOR LOADING MATRIX	150
7.4.1	<i>Factor descriptions</i>	152
7.4.2	<i>Spatial distribution of factors</i>	155
7.5	CONCLUSION	175

CHAPTER 8 RESULTS OF STAGE 2: RELATING THE RLS TO THE MARKETPLACE.....	177
8.1 INTRODUCTION.....	177
8.2 RESULTS.....	179
8.2.1 <i>Step 1: Preliminary analysis</i>	179
8.2.2 <i>Step 2: Calibration of the GWR model over the 300 metre grid</i>	186
8.2.2.1 Examining the individual parameter estimates.....	188
8.2.3 <i>Step 3: Derivation of geospatial submarket boundaries</i>	203
8.2.3.1 Error variability measures.....	204
8.2.3.2 Understanding the structure of the geospatial submarkets.....	211
8.2.3.3 User criteria.....	222
8.3 CONCLUSION.....	224
CHAPTER 9 CONCLUSION AND FUTURE RESEARCH.....	227
9.1 INTRODUCTION.....	227
9.2 MEETING THE OBJECTIVES.....	229
9.3 STAGE 1.....	232
9.3.1 <i>Weaknesses</i>	232
9.3.2 <i>Strengths</i>	234
9.4 STAGE 2.....	235
9.4.1 <i>Weaknesses</i>	235
9.4.2 <i>Strengths</i>	236
9.5 MEETING THE NEEDS.....	237
9.6 FUTURE RESEARCH.....	238
9.6.1 <i>Temporal aspects of geographic submarkets</i>	239
9.6.2 <i>Relationship of nested structural submarkets within the geospatial submarkets</i>	240
9.6.3 <i>Data</i>	240
9.6.4 <i>Identification of a wider user base</i>	242
9.7 CONCLUSION.....	242
APPENDICES.....	244
APPENDIX 7.1 - TOTAL VARIANCE EXPLAINED.....	245
APPENDIX 7.2 - ROTATED COMPONENT MATRIX.....	246
REFERENCES	247

LIST OF TABLES

TABLE 2.1	THE USE OF SUBMARKETS	27
TABLE 3.1	POPULATIONS OF AUSTRALIAN STATES AND THEIR RESPECTIVE CAPITAL CITIES, 2001.	57
TABLE 3.2	VARIABLES USED IN THE STUDY FOR STAGE 1	84
TABLE 4.1	GROUPED STUDIES AND ANALYTICAL TECHNIQUES	93
TABLE 7.1	COMPARISON OF COMMUNALITIES.....	147
TABLE 7.2	COMPONENT CORRELATION MATRIX USING THE OBLIQUE ROTATION – 10 FACTOR SOLUTION	149
TABLE 7.3	DESCRIPTION OF THE 10 FACTORS COMPRISING THE UNDERLYING DIMENSIONS OF THE RLS .	151
TABLE 8.1	MONTE CARLO RESULTS	180
TABLE 8.2	GWR MODEL SUMMARIES	182
TABLE 8.3	GLOBAL T-VALUES (MODEL 2 IN TABLE 8.2).....	186
TABLE 8.4	SUMMARY OF THE PARAMETER ESTIMATE VALUES FROM THE GWR MODEL 2 DESCRIBED IN TABLE 8.2	188
TABLE 8.5	DETERMINING CLUSTER NUMBERS - METHODOLOGY BASED ON ‘ERROR VARIABILITY’	206

LIST OF FIGURES

FIGURE 1.1	GWR CONCEPTUAL MODEL	16
FIGURE 2.1	LOCAL HOUSING SUBMARKET STRUCTURE	39
FIGURE 2.2	SEPARATE MODELS FOR EACH SUBMARKET	41
FIGURE 2.3	BROAD COMPONENTS OF PROPERTY VALUE	49
FIGURE 3.1	LOCATION CONTEXT OF STUDY AREA	58
FIGURE 3.2	DWELLING AGE DISTRIBUTION.....	62
FIGURE 3.3	INCOME DISTRIBUTION.....	64
FIGURE 3.4	MEDIAN QUARTERLY PRICE MOVEMENTS	67
FIGURE 3.5	INTERPOLATED SALE PRICE (AUGUST 2001).....	69
FIGURE 4.1	GEOSPATIAL SUBMARKET DELINEATION PROCESS	101
FIGURE 4.2	CONCEPTUAL VIEW OF WORKING SOFTWARE ENVIRONMENT.....	107
FIGURE 5.1	THE DERIVATION OF THE RESIDENTIAL LIVING STRUCTURE (RLS)	115
FIGURE 6.1	A SPATIAL KERNEL	128
FIGURE 6.2:	DISTANCE MEASURE USED IN CLUSTER ALGORITHMS	132
FIGURE 6.3	CLUSTER SIGNATURES	133
FIGURE 6.4	DERIVATION OF GEOSPATIAL SUBMARKET ENTITIES FROM RLS	134
FIGURE 7.1	SCREE PLOT OUTPUT.....	146
FIGURE 7.2	RLS CONSTRUCT – FACTOR 1	158
FIGURE 7.3	RLS CONSTRUCT – FACTOR 2	160
FIGURE 7.4	RLS CONSTRUCT – FACTOR 4.....	162
FIGURE 7.5	RLS CONSTRUCT – FACTOR 5.....	164
FIGURE 7.6	RLS CONSTRUCT – FACTOR 6.....	166
FIGURE 7.7	RLS CONSTRUCT – FACTOR 9.....	167
FIGURE 7.8	RLS CONSTRUCT – FACTOR 7.....	170
FIGURE 7.9	RLS CONSTRUCT – FACTOR 8.....	171
FIGURE 7.10	RLS CONSTRUCT – FACTOR 3.....	173
FIGURE 7.11	RLS CONSTRUCT – FACTOR 10.....	174
FIGURE 8.1	STANDARDISED RESIDUAL PLOTS: LOCAL MODEL	
FIGURE 8.2	STANDARDISED RESIDUAL PLOTS: GLOBAL MODEL	181
FIGURE 8.3	HOT SPOT LOCATIONS	184
FIGURE 8.4	PRICE INFLUENCE OF FACTOR 1.....	189
FIGURE 8.5	PRICE INFLUENCE OF FACTOR 2.....	191
FIGURE 8.6	PRICE INFLUENCE OF FACTOR 4.....	192
FIGURE 8.7	PRICE INFLUENCE OF FACTOR 5	194
FIGURE 8.8	PRICE INFLUENCE OF FACTOR 6	195
FIGURE 8.9	PRICE INFLUENCE OF FACTOR 7	196
FIGURE 8.10	PRICE INFLUENCE OF FACTOR 8	198
FIGURE 8.11	PRICE INFLUENCE OF FACTOR 9	199
FIGURE 8.12	PRICE INFLUENCE OF FACTOR 3	201
FIGURE 8.13	PRICE INFLUENCE OF FACTOR 10	202
FIGURE 8.14	CHANGE IN AGGLOMERATION COEFFICIENT	205
FIGURE 8.15	‘10’ CLUSTER SOLUTION	
FIGURE 8.16	‘50’ CLUSTER SOLUTION.....	208
FIGURE 8.17	‘150’ CLUSTER SOLUTION	
FIGURE 8.18	‘330’ CLUSTER SOLUTION.....	209
FIGURE 8.19	MAXIMUM NUMBER OF RECOMMENDED CLUSTERS	210
FIGURE 8.20	LOCATION OF CLUSTER EXAMPLES.....	212
FIGURE 8.21	CLUSTER SIGNATURES – 10 CLUSTERS	215
FIGURE 8.22	CLUSTER SIGNATURES – EXAMPLE FROM NORTH EAST	216
FIGURE 8.23	CLUSTER SIGNATURES – EXAMPLE FROM EAST	219
FIGURE 8.24	CLUSTER SIGNATURES – EXAMPLE FROM SOUTH.....	220
FIGURE 8.25	CLUSTER SIGNATURES – EXAMPLE FROM WEST	221

ABSTRACT

While it is generally accepted that residential submarkets exist, this is not the case for either their definition or delineation. This study has developed and assessed a methodology for deriving the geospatial dimensions of residential real estate submarkets based on the behaviour of the marketplace with respect to the underlying dimensions of the residential real estate living structure (RLS). Importantly, the methodology makes no prior assumptions about where the spatial boundaries might be. They were empirically derived from the data alone. It is achieved by building on previous work in the field and seeking to identify the fundamental issues in residential real estate market behaviour. The first basic premise of the thesis is that when a dwelling is sold, the commodity traded is a piece of real estate geography comprising a complex bundle of both spatial and structural attributes. The second basic premise is the recognition in the methodology of the importance of 'location'. The price of the real estate geography varies across geographical space in a continuous fashion and it is this price variability that is defined, in this study, to be the geospatial submarket identifier. The study adopts a two-stage methodology reflecting these two basic premises.

Firstly, a complex bundle of attributes is collected for every property in the study area and distilled into its underlying dimensions using principal component analysis. The resulting factors are used in the second stage as independent variables in a hedonic geographically weighted regression model to determine the price variability across geographical space of the underlying residential real estate structure. User-defined breaks in the continuous price surface delineate the geospatial submarket boundaries.

The study represents a new approach to the delineation of geospatial submarket boundaries and is yet to be fully assessed by the two major identified users (the planning profession and the valuation profession). However, initial feedback indicates that the ability of the methodology to describe the geospatial submarket boundaries in terms of 'how' and 'where' location affects the market price of the underlying real estate geography, gives the land professional a better understanding of the submarket structure in which they are working.

DECLARATION

This thesis contains no material that has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text of the thesis.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for photocopying and loan.

Signed: _____ Date: _____

ABBREVIATIONS AND ACRONYMS

ABS	Australian Bureau of Statistics
AIC	Akaike Information Criteria statistic
CA	Cluster Analysis
CAMA	Computer Assisted Mass Appraisal
CBD	Central Business District
CD	Census Collection District (ABS definition)
DCDB	Digital Cadastral Data Base
GDP	Gross Domestic Product
GIS	Geographic Information System
GISCA	The National Centre for Social Applications of Geographical Information Systems, The University of Adelaide.
GWR	Geographically Weighted Regression
HI	Hedonic Index
HM	Homogeneity Measure
IDW	Inverse Distance Weighted (a data interpolation methodology)
LIS	Land Information System
LOTS	Land Ownership and Tenure System
LVS	Location Value Signature
MRA	Multiple Regression Analysis
NDVI	Normalised Density of Vegetation Index
PCA	Principal Component Analysis
RLS	Residential Living Structure
VG	Valuer General

ACKNOWLEDGEMENTS

This study was undertaken on an Australian Postgraduate Award (The University of Adelaide) with support from the Department of Geographical and Environmental Studies.

My thanks are due to many people whose help varied from casual comments to formal constructive feedback, all of which contributed to a better understanding of the problems associated with an undertaking of this nature. Particular thanks go to Professor Graeme Hugo, my principal supervisor, whose continued guidance, encouragement and insights were invaluable. Thanks to my co-supervisor, Dr Kevin Harris, and to Dr. Jack Massey, whose readings of early drafts were incisive and encouraging. After Kevin's retirement, sincere thanks to Associate Professor Clive Forster, who took over the role as co-supervisor, providing clear and constructive guidance at a critical time in the candidature. Also thanks to Dr. Bertram Ostendorf (The University of Adelaide), Graeme Tucker (South Australian Department of Health) and Dr. Mark Lethbridge (Flinders University) who were generous in willingly denoting their valuable time to advising on various statistical aspects of this study.

Special thanks to the ARC Research Network in Spatially Integrated Social Science for their Summer School held over a 10-day period in Brisbane, in which students in various stages of their candidature were brought together to share experiences and receive invaluable instruction and encouragement in a wide range of aspects concerned with the completion of a piece of academic social science research. This was further supported by the running of specific events, including a 4-day workshop in Adelaide in the use of geographically weighted regression conducted by Professor Stewart Fotheringham and

Martin Charlton from the National Centre for Geocomputation, National University of Ireland, Maynooth. This was an essential element in the completion of the study and much appreciated.

My thanks to the National Centre for the Social Applications of GIS (GISCA) for their support, not only in terms of infrastructure, but also GIS expertise. In particular, Neil Coffee, whose patience and professional guidance through the maze of GIS idiosyncrasies was invaluable. Also, particular thanks to David Coombe for his unique ability to solve the inevitable problems associated with the IT environment so important in managing the large amounts of spatial data involved in this study.

Thanks to my fellow students for creating an excellent, and fun, working environment in which to undertake academic research, especially to Natasha Howard, who in the hectic days of collating the final product, provided much needed support by willingly sharing her valuable word processing skills.