

Using Aggregated Demographic Data To  
Inform Electoral Boundary  
Redistributions: 2010 South Australian  
Election

Casey Briggs

*Thesis submitted for the degree of  
Master of Philosophy  
in  
Statistics and Applied Mathematics  
at  
The University of Adelaide*

School of Mathematical Sciences

April 2015

# Contents

<b>Signed Statement</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>viii</b>
<b>Abstract</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Background . . . . .	4
1.3 Thesis Outline . . . . .	5
<b>2 Literature Review and Background</b>	<b>9</b>
2.1 South Australia’s electoral system . . . . .	9
2.2 South Australian electoral history . . . . .	12
2.3 Electoral Districts Boundaries Commission . . . . .	17
2.3.1 Current rules for redistributions . . . . .	19

2.4	The fairness clause in operation . . . . .	21
2.4.1	Methodology of the EDBC . . . . .	23
2.5	Collection Districts . . . . .	26
2.6	Factors that influence voting intention . . . . .	27
2.7	Discussion . . . . .	30
<b>3</b>	<b>Data</b>	<b>32</b>
3.1	Raw Data . . . . .	32
3.2	Electoral Data . . . . .	36
3.2.1	Electoral District Results . . . . .	36
3.2.2	Polling Place Results . . . . .	37
3.2.3	Voter Location Data . . . . .	37
3.3	Census Data . . . . .	38
3.3.1	The ‘Other’ Category . . . . .	42
3.4	Exploring Collection Districts . . . . .	43
3.4.1	Distributions of predictors . . . . .	43
3.4.2	Aggregate Statistics . . . . .	44
3.5	Visualising Electoral Districts . . . . .	50
3.6	Extrapolation . . . . .	55
3.7	Correlation . . . . .	56

3.8	Principal Component Analysis . . . . .	57
3.8.1	PCA applied . . . . .	58
3.8.2	Clusters . . . . .	63
3.8.3	Structure within PCA . . . . .	65
3.9	Conclusion . . . . .	69
<b>4</b>	<b>Modelling at Electoral District Level</b>	<b>70</b>
4.1	Statistical Background . . . . .	72
4.1.1	Multiple Linear Regression . . . . .	72
4.1.2	Multiple Logistic Regression . . . . .	73
4.1.3	Multinomial Logistic Regression . . . . .	74
4.2	Method of Model Validation . . . . .	75
4.3	Model 4.1 - with original 17 predictors . . . . .	85
4.4	Model 4.2 - with all principal components . . . . .	90
4.5	Model 4.3 - with two principal components . . . . .	91
4.6	Discussion . . . . .	95
<b>5</b>	<b>Modelling at Polling Place Level</b>	<b>97</b>
5.1	Voter Location Data . . . . .	98
5.1.1	Method of Voter Location Data Generation . . . . .	98
5.1.2	Data Verification and Cleaning . . . . .	100

5.2	Demographics of Polling Places . . . . .	110
5.3	Model 5.1 - with original 17 predictors . . . . .	111
5.4	Model 5.2 - with two principal components . . . . .	116
5.5	Comparing Multinomial Regression Models . . . . .	120
5.5.1	Summary Statistics . . . . .	121
5.5.2	Comparing ‘goodness of fit’ . . . . .	122
5.5.3	Graphical Analysis of Fit . . . . .	124
5.5.4	Preferred Model . . . . .	126
5.6	Discussion . . . . .	127
<b>6</b>	<b>Spatially-aware Model</b>	<b>128</b>
6.1	Model Description . . . . .	130
6.1.1	Implementation . . . . .	133
6.2	Model 6.1 - Electoral district level . . . . .	133
6.3	Model 6.2 - Polling place level . . . . .	137
6.4	Model 6.3 - logistic analogue of Model 5.1 . . . . .	142
6.5	Model Predictions . . . . .	146
6.5.1	Goodness of fit . . . . .	150
6.6	Preferred Model . . . . .	153
6.7	Coefficients . . . . .	154

6.8	Discussion . . . . .	155
<b>7</b>	<b>Comparing to EDBC predictions</b>	<b>157</b>
<b>8</b>	<b>Conclusion</b>	<b>166</b>
8.1	Ideas for future research . . . . .	167
<b>A</b>	<b>Supplementary Material for Chapters 4 &amp; 5</b>	<b>170</b>
A.1	Mismatches in Voter Location Data . . . . .	171
A.2	Multinomial Regression Models . . . . .	172
A.2.1	Model 4.1 . . . . .	172
A.2.2	Model 4.2 . . . . .	174
A.2.3	Model 4.3 . . . . .	175
A.2.4	Model 5.1 . . . . .	177
A.2.5	Model 5.2 . . . . .	179
A.3	Model 4.2 Validation . . . . .	180
A.4	Side by side views of multinomial graphics . . . . .	183
<b>B</b>	<b>Supplementary Material for Chapter 6</b>	<b>189</b>
B.1	Full Model 6.2 Output . . . . .	190
B.2	Full Model 6.3 Output . . . . .	191
	<b>Bibliography</b>	<b>192</b>



# Signed Statement

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name 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. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

SIGNED: ..... DATE: .....



# Acknowledgements

To my supervisors, Professor Nigel Bean, Dr Jonathan Tuke, and Professor Clement Macintyre: this thesis wouldn't exist without your support. Thank you for persisting with me and being so generous with your time and expertise. Thanks also to my former supervisor, Dr David Green.

Thank you to David Gully, Deputy Electoral Commissioner at the Electoral Commission of South Australia, for supplying data for this thesis, and responding helpfully to my questions even in the midst of busy election periods. My appreciation also to everyone that read drafts and provided feedback, in particular Sam and Stella who read the complete draft.

I want to thank my family, and in particular my parents, for supporting me through my entire education. There is no way I would be in the position I am now without your encouragement and love, and I'm deeply grateful to you for the opportunities you've given me.

Thanks to all the staff and students of the School of Mathematical Sciences, of whom there are far too many to name, for creating a great place to work since I commenced as an undergraduate in 2008.

# Abstract

Electoral boundaries in South Australia are currently a contentious issue in politics, with allegations that the current boundaries are unfair. South Australia has fairness provisions that are unique in Australia governing the boundaries of electoral districts. However, in three of the last six state elections, the objective of fairness as characterised by these provisions has not been met.

Boundaries are drawn by the independent Electoral Districts Boundaries Commission, and are revised after every general election in South Australia. The Commission's method uses estimates for the voting behaviours in small areas to inform the decisions about boundary changes.

The objective of this thesis is to develop an alternative method for calculating these estimates, and test the credibility of the resultant estimates from our new method.

We develop a series of gradually refined regression models that use demographic data in South Australia to predict voting behaviour. The demographic data is sourced from the periodical Census of Population and Housing. In this research we also test the proposition that income, education level, and the language people speak at home are significant factors in their voting behaviour, at an aggregated group level.

We contend that the predictions calculated under the preferred model in this thesis are credible, and that the techniques used warrant further exploration.