Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.
A SYSTEMATIC INVESTIGATION OF THE ESTIMATION OF THE DIRICHLET MODEL

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Marketing at Massey University

Zane Kearns

1999
ABSTRACT

The NBD/Dirichlet is a stochastic model of purchase incidence and brand choice which parsimoniously integrates a wide range of well-established empirical regularities in fast moving consumer goods markets. More recently this work has been extended into other areas such as the prescribing of pharmaceuticals, (Stern 1994); airline aviation fuel contracts, (Uncles and Ehrenberg 1990a); and the visiting of retail store chains, (Uncles and Ehrenberg 1990b).

By combining the stochastic assumptions of the model, namely Poisson purchasing of products, with mean rate distributed gamma across the population, and brand choice represented by multinomial probabilities distributed Dirichlet across consumers; a number of aspects of the aggregate behaviour of consumers can be successfully predicted successfully.

This thesis examines the estimation issues in the Dirichlet model, specifically, the central Dirichlet parameter \( S \) used to represent heterogeneity in brand choice.
ACKNOWLEDGEMENTS

In the first instance I would like to thank Professors A S C Ehrenberg, G J Goodhardt, and C Chatfield for their development of what has become known as the Dirichlet model. It is difficult to discern the unique contributions of these researchers, but overall their efforts represent a serious but often unrecognised contribution to marketing and the study of consumer behaviour.

I would also like to thank Professor Ehrenberg more personally for the time taken to answer my queries about the model, and being overly kind in not pointing out how silly some of my questions were, or that I should really have known the answer.

Of course the greatest burden of gratitude goes to my supervisor Associate Professor A C Lewis. His guidance is evident throughout this thesis, and without it, the thesis would be undeniably poorer. Thanks must also go to Dr Greg Arnold who lent valuable advice.
# TABLE OF CONTENTS

## Chapter One

Introduction

1.1 Background 1
1.2 Objectives of the Thesis 2
1.3 Outline of the Chapters 4

## Chapter Two

Stochastic Modelling of Consumer Behaviour

2.1 Introduction 5
2.2 Loyalty 5
2.3 Stochastic Modelling of Consumer Behaviour 8
2.4 The Dirichlet model of Consumer Behaviour 10
   2.4.1 Aggregative Considerations 10
2.5 Dirichlet Findings 12
   2.5.1 The Fit of the Dirichlet model 15
   2.5.2 Known Systematic Discrepancies 24
2.6 Summary 25

## Chapter Three

Assumptions of the Dirichlet model

3.1 Introduction 26
3.2 The Purchase Incidence Model 27
   3.2.1 The Poisson Process 27
   3.2.2 Gamma Heterogeneity 32
   3.2.3 Negative Binomial Distribution of Purchases 35
3.3 The Brand Choice Model 39
   3.3.1 Multinomial Process 39
   3.3.2 Dirichlet Heterogeneity of Choice Probabilities 44
3.4 Summary 46
Chapter Four  
Estimation of the Dirichlet model

4.1 Introduction  
4.2 Background  
4.3 Product Class Purchasing  
4.3.1 Estimation  
4.3.2 Comparisons  
4.4 Brand Choice  
4.4.1 Estimation  
4.4.2 Alternatives and Comparisons  
4.5 Summary

Chapter Five  
Methodology

5.1 Introduction  
5.2 Method

Chapter Six  
Results

6.1 Introduction  
6.2 Parameters of the Sampling Distributions of the Simulated Panels  
6.3 Estimates of the Dirichlet $S$ Parameter  
6.4 Effects of Sampling Error in the Dirichlet Predictions

Chapter Seven  
Discussion

7.1 Introduction  
7.2 Background  
7.3 Results

References
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Brand performance measures for Laundry Detergent</td>
<td>15</td>
</tr>
<tr>
<td>2.2</td>
<td>Conditions under which the Dirichlet patterns are known to occur</td>
<td>17</td>
</tr>
<tr>
<td>2.3</td>
<td>Annual performance measures for the eight leading brands</td>
<td>19</td>
</tr>
<tr>
<td>2.4</td>
<td>Predictions of purchase frequency and share of requirements</td>
<td>20</td>
</tr>
<tr>
<td>2.5</td>
<td>Market share and share of requirements</td>
<td>23</td>
</tr>
<tr>
<td>3.1</td>
<td>A stochastic representation of consumer purchases (x)</td>
<td>27</td>
</tr>
<tr>
<td>3.2</td>
<td>Purchase incidence models and their properties</td>
<td>38</td>
</tr>
<tr>
<td>3.3</td>
<td>A stochastic representation of brand choice</td>
<td>39</td>
</tr>
<tr>
<td>3.4</td>
<td>Summary studies on the order of the brand choice process</td>
<td>41</td>
</tr>
<tr>
<td>4.1</td>
<td>&quot;Constant D&quot; approximation predictions of duplication and Bemmaor's estimator</td>
<td>57</td>
</tr>
<tr>
<td>5.1</td>
<td>Simulation parameters</td>
<td>62</td>
</tr>
<tr>
<td>6.1</td>
<td>Recovery of simulation parameters</td>
<td>64</td>
</tr>
<tr>
<td>6.2</td>
<td>Comparison of the standard error of $m$ obtained from the simulated samples and theoretical value.</td>
<td>65</td>
</tr>
<tr>
<td>6.3</td>
<td>Standard Deviations of the Estimates of Penetration and Frequency from the Simulated Samples</td>
<td>67</td>
</tr>
<tr>
<td>6.4</td>
<td>Kurtosis, Skewness and Kolmogorov-Smirnov Statistics from the Simulated Samples</td>
<td>68</td>
</tr>
<tr>
<td>6.5</td>
<td>Mean and Range of Estimates of $S$</td>
<td>71</td>
</tr>
<tr>
<td>6.6</td>
<td>Ratio of Range in $S(n)$ to $S(1000)$</td>
<td>72</td>
</tr>
<tr>
<td>6.7</td>
<td>Ratio of range in estimates of $S$ to the standard error of the mean rate of buying</td>
<td>74</td>
</tr>
<tr>
<td>6.8</td>
<td>Ratios of the Standard error of the mean rate of buying to range of Dirichlet Predictions: Simulation A</td>
<td>76</td>
</tr>
<tr>
<td>6.9</td>
<td>Ratios of the Standard error of the mean rate of buying to range of Dirichlet Predictions: Simulation B</td>
<td>77</td>
</tr>
<tr>
<td>6.10</td>
<td>Ratios of the Standard error of the mean rate of buying to range of Dirichlet Predictions: Simulation C</td>
<td>78</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>6.1</th>
<th>Histograms and Normal Probability Plots for Brand C Parameters</th>
<th>69</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2</td>
<td>Simulated values of penetration and frequency and resulting estimates of $S$</td>
<td>73</td>
</tr>
</tbody>
</table>