

SPECIFICATION AND ESTIMATION OF STOCHASTIC FRONTIER PRODUCTION FUNCTIONS

by

Timothy James Coelli

B.App.Ec., Dip.Comp.Sci. (NE)

**A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF
PHILOSOPHY OF THE UNIVERSITY OF NEW ENGLAND**

January 1996

ACKNOWLEDGMENTS

First and foremost I must thank my supervisor, Associate Professor George Battese. From October 1984, when George first introduced me to stochastic frontier production functions, to the date of submission of this thesis, George has had a significant influence upon much of my research, and upon this thesis in particular. This is particularly evident in that Chapters 3 to 5 of this thesis are based upon papers co-authored with George. Thank you George for your leadership, encouragement and good humour.

I would also like to thank my other colleagues in the Department of Econometrics. In particular, I have significantly benefited from discussions with Professor Bill Griffiths, Associate Professor Prasada Rao and Associate Professor Howard Doran.

I would like to acknowledge the financial support of the Australian Electricity Supply Industry Research Board which partially funded the study discussed in Chapter 6.

I wish to thank a large number of people for comments they have provided on various papers associated with this thesis. In particular, I thank Tom Cowing, Gary Ferrier, Bill Greene, Carter Hill, Bill Malcolm and seminar audiences at the University of New England, Louisiana State University and University of Arkansas for comments. I also wish to thank Sergio Perelman, Knox Lovell and many other users of my FRONTIER computer program for providing feedback on its various versions. Furthermore, I would like to thank conference audiences at a number of Australasian Meetings of the Econometric Society and the Annual Conferences of the Australian Agricultural Economics Society for their many useful comments. I also wish to acknowledge the many constructive comments provided by anonymous referees from the *Journal of Productivity Analysis*, *Journal of Econometrics*, *Journal of Empirical Economics*, *Australian Journal of Agricultural Economics*, *American Journal of Agricultural Economics*, and *Economics Letters*.

Finally, and most importantly, I would like to thank my family. To my wife, Michelle, and daughters, Kathleen and Emily, thank you for your support and understanding over the past four years when I have foolishly spent more time with my PC than with you. I am also greatly indebted to my parents who worked very hard to ensure that

myself and my six brothers and sisters have had the opportunity to obtain a quality education. This thesis is dedicated to my father, John Kevin Coelli, who passed away in 1992.

ABSTRACT

This thesis is concerned with the specification, estimation, application and testing of stochastic frontier production functions. The majority of the material in this document is derived from seven research papers written over the past four years. The thesis begins with a review of frontier estimation methods, with an emphasis on information relevant to agricultural applications. The two primary methods of frontier estimation, stochastic frontiers and data envelopment analysis (DEA), are described and compared. The influence of data noise is observed to be crucial in the selection of appropriate methods.

A stochastic frontier production function for panel data in which the technical inefficiency effects are specified to be the product of a deterministic exponential function of time and time-invariant inefficiency effects is specified. This model specification is introduced to address criticisms of earlier stochastic frontier specifications for panel data which assume time-invariant technical inefficiency effects. An application involving Indian paddy farmers illustrates the use of the model.

A stochastic frontier production model for panel data in which the technical inefficiency effects are permitted to be a function of firm-specific variables and time is also specified. This model is an improvement over the two-stage approach to the modelling of technical inefficiency effects, which is inconsistent in its statistical assumptions. An empirical application using the above mentioned data, along with information on the age and formal schooling of the farmers, indicates that age and education do have a significant influence upon the levels of the technical inefficiency effects of these Indian paddy farmers. Two additional applications of this model specification are also conducted. The first of these involves a larger panel data set on farmers from three Indian villages. In this application, farmer age and years of formal schooling, along with farm size, are found to influence the technical inefficiency effects of the farmers in two of the three villages considered. The second application involves panel data on electricity generation in Australian coal-fired power plants. In this application, capacity factor, unit size, plant vintage and coal quality are shown to have a significant influence upon plant inefficiency.

Following this, the computer program, FRONTIER Version 4.1, which has been written to estimate the above-mentioned frontier models, is described in detail. This program can also be used to estimate a number of other model specifications, including cross-sectional models and cost frontiers. The program calculates maximum-likelihood estimates, their estimated standard errors, and technical efficiency predictions.

Finally, Monte Carlo methods are used to investigate the finite-sample properties of estimators for parameters of a stochastic frontier production function in which the technical inefficiency effects have half-normal distribution. The relative performance of the maximum-likelihood (ML) and corrected ordinary least-squares (COLS) estimators are investigated, together with five alternative test statistics for testing for the absence of technical inefficiency effects. The results indicate that there is substantial bias in both ML and COLS estimators when the percentage contribution of inefficiency in the composed error term is small. It is also concluded that ML estimators should be used in preference to COLS estimators because of large mean squared error advantages when the relative contribution of the variance of the technical inefficiency effects is greater than 0.5. The results also show that Wald and likelihood-ratio tests of the null hypothesis of no technical inefficiency effects have incorrect size, and that a one-sided likelihood-ratio test and a test of the third-moment of the OLS residuals have correct size, with the one-sided likelihood-ratio test having the better power of the two.

CONTENTS

	Page
ACKNOWLEDGMENTS	iii
ABSTRACT	v
LIST OF TABLES.....	x
LIST OF FIGURES.....	xii
1. INTRODUCTION.....	1
1.1 FRONTIER FUNCTIONS AND EFFICIENCY MEASUREMENT	1
1.2 OUTLINE OF THE THESIS	2
2. LITERATURE REVIEW	5
2.1 INTRODUCTION.....	5
2.2 EARLY LITERATURE.....	7
2.3 STOCHASTIC FRONTIERS	11
2.3.1 <i>Stochastic Frontier Production Functions</i>	11
2.3.2 <i>Estimation Methods</i>	12
2.3.3 <i>Alternative Functional Forms</i>	12
2.3.4 <i>Dual Forms of the Technology</i>	13
2.3.5 <i>Panel Data</i>	16
2.3.6 <i>Determinants of Inefficiency</i>	18
2.4 DATA ENVELOPMENT ANALYSIS.....	19
2.4.1 <i>The Constant Returns-to-scale (CRS) Model</i>	20
2.4.2 <i>The Variable Returns-to-scale (VRS) Model</i>	24
2.4.3 <i>Output-orientated Models</i>	25
2.4.4 <i>Other Variants and Extensions</i>	26
2.5 APPLICATIONS TO AGRICULTURE.....	27
2.6 CONCLUSIONS	29
3. A STOCHASTIC FRONTIER PRODUCTION FUNCTION WITH TIME- VARYING INEFFICIENCY EFFECTS.....	32
3.1 INTRODUCTION.....	32
3.2 MODEL SPECIFICATION	33
3.3 EMPIRICAL EXAMPLE.....	36
3.4 CONCLUSIONS	47
4. A STOCHASTIC FRONTIER PRODUCTION FUNCTION INCORPORATING A MODEL FOR TECHNICAL INEFFICIENCY EFFECTS.....	49
4.1 INTRODUCTION.....	49
4.2 MODEL SPECIFICATION	51
4.3 EMPIRICAL EXAMPLE.....	54
4.4 CONCLUSIONS	61
5. IDENTIFICATION OF FACTORS WHICH INFLUENCE THE TECHNICAL INEFFICIENCY OF INDIAN FARMERS.....	63

5.1 INTRODUCTION.....	63
5.2 PANEL DATA ON INDIAN AGRICULTURE	65
5.3 THE STOCHASTIC FRONTIER AND INEFFICIENCY MODEL	67
5.4 RESULTS AND DISCUSSION.....	69
5.4.1 <i>Maximum-likelihood Estimates</i>	69
5.4.2 <i>Tests of Hypotheses</i>	74
5.4.3 <i>Technical Efficiencies of Farmers</i>	77
5.5 CONCLUSIONS	84
6. MEASUREMENT AND SOURCES OF TECHNICAL INEFFICIENCY IN AUSTRALIAN COAL-FIRED ELECTRICITY GENERATION.....	86
6.1 INTRODUCTION.....	86
6.2 LITERATURE	88
6.2.1 <i>Non-frontier Analyses</i>	89
6.2.2 <i>Frontier Analyses</i>	90
6.3 DATA AND MODEL SPECIFICATION.....	91
6.3.1 <i>Data</i>	91
6.3.2 <i>Model Specification</i>	93
6.4 RESULTS AND DISCUSSION.....	96
6.4.1 <i>Maximum-likelihood Estimates</i>	96
6.4.2 <i>Tests of Hypotheses</i>	98
6.4.3 <i>Economic Plausibility of the Results</i>	100
6.4.4 <i>A Comparison with the Two-stage Approach</i>	108
6.5 CONCLUSIONS	109
7. A COMPUTER PROGRAM FOR ESTIMATION OF STOCHASTIC FRONTIER PRODUCTION AND COST FUNCTIONS: FRONTIER VERSION 4.1	111
7.1 INTRODUCTION.....	111
7.2 MODEL SPECIFICATIONS	112
7.2.1 <i>Model 1: The Battese and Coelli (1992) Specification</i>	113
7.2.2 <i>Model 2: The Battese and Coelli (1995) Specification</i>	115
7.2.3 <i>Cost Functions</i>	116
7.2.4 <i>Efficiency Predictions</i>	118
7.3 THE FRONTIER PROGRAM.....	119
7.3.1 <i>Files Needed</i>	119
7.3.2 <i>The Three-step Estimation Method</i>	121
7.3.3 <i>Program Output</i>	122
7.3.4 <i>Differences Between Versions 2.0 and 4.1</i>	123
7.4 SOME SHORT EXAMPLES.....	126
7.4.1 <i>A Cobb-Douglas Production Frontier using Cross-sectional Data and Assuming a Half-Normal Distribution</i>	126
7.4.2 <i>A Translog Production Frontier using Cross-sectional Data and Assuming a Truncated Normal Distribution</i>	131
7.4.3 <i>A Cobb-Douglas Cost Frontier using Cross-sectional Data and Assuming a Half-Normal Distribution</i>	134
7.4.4 <i>The Battese and Coelli (1992) Specification (Model 1)</i>	136
7.4.5 <i>The Battese and Coelli (1995) Specification (Model 2)</i>	139
7.5 FINAL POINTS	141

8. A MONTE CARLO ANALYSIS OF A STOCHASTIC FRONTIER PRODUCTION FUNCTION.....	142
8.1 INTRODUCTION.....	142
8.2 METHODS OF ESTIMATION	144
8.3 TESTS OF HYPOTHESES	148
8.4 DESIGN OF THE MONTE CARLO EXPERIMENT	152
8.5 MONTE CARLO RESULTS.....	155
8.5.1 <i>ML and COLS Estimators</i>	155
8.5.2 <i>Mean Technical Efficiency</i>	162
8.5.3 <i>Variances of ML and COLS Estimators</i>	164
8.5.4 <i>Tests of Hypotheses</i>	164
8.6 CONCLUSIONS	170
9. CONCLUSIONS.....	172
9.1 SUMMARY OF RESULTS.....	172
9.2 FUTURE WORK	174
APPENDICES.....	177
APPENDIX 1: DERIVATIONS FOR THE TIME-VARYING INEFFICIENCY MODEL.....	177
APPENDIX 2: DERIVATIONS FOR THE STOCHASTIC FRONTIER AND INEFFICIENCY MODEL.....	182
APPENDIX 3: ELECTRICITY DATA	186
APPENDIX 4: FRONTIER PROGRAMMER'S GUIDE.....	191
APPENDIX 5: FRONTIER PROGRAM LISTING	195
APPENDIX 6: ASYMPTOTIC STANDARD ERRORS OF THE COLS ESTIMATORS	222
APPENDIX 7: SHAZAM CODE FOR COLS.....	225
REFERENCES	226

LIST OF TABLES

Table	page
2.1 Applications of Frontier Models to Agriculture, 1985-1996	28
3.1 Summary Statistics for Variables in the Stochastic Frontier Production Function for Paddy Farmers in Aurepalle	38
3.2 Maximum-likelihood Estimates for Parameters of Stochastic Frontier Production Functions for Aurepalle Paddy Farmers	41
3.3 Tests of Hypotheses for Parameters of the Distribution of the Inefficiency Effects	42
3.4 Predicted Efficiencies of Paddy Farmers in Aurepalle for the years, 1975-76 through 1984-85	43
3.5 Maximum-likelihood Estimates for Parameters of Production Functions Which Account for Technical Change	45
4.1 Maximum-likelihood Estimates for Parameters of Stochastic Frontier Production Functions and Inefficiency Models for Paddy Farmers in Aurepalle	57
4.2 Tests of Hypotheses for Parameters of the Stochastic Frontier and Inefficiency Model for Paddy Farmers in Aurepalle	59
4.3 Technical Efficiencies of Paddy Farmers in Aurepalle	61
5.1 Summary Statistics for Variables in the Stochastic Frontier and Inefficiency Models for Farmers in Three Indian Villages	70
5.2 Maximum-likelihood Estimates for Parameters of the Stochastic Frontier and Inefficiency Models for Three Indian Villages	72
5.3 Tests of Hypotheses for Coefficients of the Explanatory Variables for the Technical Inefficiency Effects in Stochastic Frontier Production Functions for Three Indian Villages	75
5.4 Statistics for Tests of Hypotheses Involving Some Coefficients of the Stochastic Frontier Production Functions for Three Indian Villages	77
5.5 Predicted Technical Efficiencies for Farmers in Aurepalle	79
5.6 Predicted Technical Efficiencies for Farmers in Kanzara	80
5.7 Predicted Technical Efficiencies for Farmers in Shirapur	81
6.1 Summary Statistics for Observations on 13 Coal-fired Electricity Generating Plants in Australia during 1981-82 to 1990-91	94
6.2 Maximum-likelihood Estimates of the Stochastic Frontier and Inefficiency Model for Electricity Generation in Australia	97
6.3 Tests of Hypotheses of Parameters of the Stochastic Frontier and Inefficiency Model for Electricity Generation in Australia	99
6.4 Key Estimates Derived From the Translog Frontier and Inefficiency Model With Hicks-Neutral Technical Change	101

6.5	Technical Efficiencies for 13 Power Plants in Australia, 1981-82 to 1990-91	105
6.6	OLS Estimates of the Second-stage Regression of Technical Inefficiency Effects	110
7.1a	Listing of Data File, EG1.DAT	128
7.1b	Listing of Shazam Instruction File, EG1.SHA	128
7.1c	Listing of Data File, EG1.DTA	129
7.1d	Listing of Instruction File, EG1.INS	129
7.1e	Listing of Output File, EG1.OUT	130
7.2a	Listing of Data File, EG2.DAT	132
7.2b	Listing of Shazam Instruction File, EG2.SHA	132
7.2c	Listing of Data File, EG2.DTA	133
7.2d	Listing of Instruction File, EG2.INS	133
7.3a	Listing of Data File, EG3.DAT	134
7.3b	Listing of Shazam Instruction File, EG3.SHA	134
7.3c	Listing of Data File, EG3.DTA	135
7.3d	Listing of Instruction File, EG3.INS	135
7.4a	Listing of Data File, EG4.DAT	136
7.4b	Listing of Shazam Instruction File, EG4.SHA	137
7.4c	Listing of Data File, EG4.DTA	138
7.4d	Listing of Instruction File, EG4.INS	138
7.5a	Listing of Data File, EG5.DAT	139
7.5b	Listing of Shazam Instruction File, EG5.SHA	139
7.5c	Listing of Data File, EG5.DTA	140
7.5d	Listing of Instruction File, EG5.INS	140
8.1	Bias, Variance and MSE for ML Estimators	156
8.2	Bias, Variance and MSE for COLS Estimators	159
8.3	Bias, Variance and MSE in ML Estimator of Mean Technical Efficiency	163
8.4	Variances of ML Estimators	165
8.5	Variances of COLS Estimators	166
8.6	Percentage of Rejections of the Null Hypothesis that Gamma is Zero	168
A3.1	Energy Contents of Fuels (in megajoules/kilogram) in 13 Power Plants in Australia	187
A3.2	Data From 13 Australian Coal-Fired Power Plants	187
A4.1	The Start-up File for the FRONTIER Program: FRONT41.000	191

LIST OF FIGURES

Figure	page
2.1 Technical and Allocative Efficiencies	8
2.2 Piecewise-linear Convex Isoquant	9
2.3 Efficiency Measurement and Input Slacks	23
2.4 Output- and Input-based Efficiency Measures	26
3.1 Predicted Technical Efficiencies of Aurepalle Paddy Farmers	44
5.1 Frequency Distribution of Predicted Technical Efficiencies of Farmers in Aurepalle	82
5.2 Frequency Distribution of Predicted Technical Efficiencies of Farmers in Kanzara	82
5.3 Frequency Distribution of Predicted Technical Efficiencies of Farmers in Shirapur	83
5.4 Mean Technical Efficiencies of the Three Indian Villages	83
6.1a Technical Efficiencies for 13 Power Plants in Australia, 1981-82 to 1990-91	106
6.1b Technical Efficiencies for 13 Power Plants in Australia, 1981-82 to 1990-91 (enlargement)	106
6.2 Mean Technical Efficiencies for 13 Power Plants in Australia, 1981-82 to 1990-91	107
6.3 Annual Mean Technical Efficiencies of 13 Power Plants in Australia, 1981-82 to 1990-91	107
7.1 Truncated Normal Densities	125
8.1 Bias in the ML Estimator of γ	158
8.2 MSE of the ML Estimator of γ	158
8.3 Bias in the ML and COLS Estimators of γ	160
8.4 Ratio of COLS MSE to ML MSE for γ	160
8.5 Power of Tests of Hypotheses when the Sample Size is 100	169
8.6 Power of the One-sided Likelihood-ratio Test	169