

The University of Adelaide

**Semiparametric Models with
Endogeneity and their Application to an
Empirical Demand Analysis**

by

Nam-Hyun Kim

This thesis is presented for the degree of

Doctor of Philosophy

at

the School of Economics

April 2013

For a thesis that does not contain work already in the public domain

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.

Signature

Date 23/08/2013

Acknowledgements

I would like to take this opportunity to express my sincere gratitude to a number of people and institutions, who have provided me with useful assistance and support at various stages of my PhD study.

Firstly, I thank my supervisors: Professor Jiti Gao, Dr Duygu Yengin and Dr Nicholas Sim. In specific, I offer my sincerest gratitude to Professor Jiti Gao, who has supported me throughout my thesis with his patience and knowledge. I also thank a number of faculty members at the School of Economics, The University of Adelaide, for their assistance, both while I was there and when I was a visiting PhD student at other institutions. Special thanks go to Professor Christopher Findlay, Dr Mandar Oak, Dr Ralph Bayer and Dr Jacob Wong for their assistance with administrative issues on various occasions. In addition, I am sincerely grateful to Dr Seungmoon Choi for his kind assistance and support on academic and various other issues, which helped me to overcome the difficult obstacles. I also thank Professor Fabrice Collard at the Department of Economics, The University of Bern, for his kind assistance with my presentation skills and his support.

I also thank the Department of Econometrics and Business Statistics, Monash University, especially Professor Brett Inder, for the opportunity to visit the department.

Similarly, many thanks to all the faculty members at the Department of Mathematics and Statistics, The University of Canterbury, for giving me such a warm welcome and hospitality, especially Professor Jennifer Brown, Dr Marco Reale, Dr Carl Scarrott, Dr Qui Bui and Dr Dominic Lee. Furthermore, Dr Marco Reale has broadened my research scope and perspective, and Dr Dominic Lee has provided me with such invaluable suggestions on the application of the work in this thesis.

This thesis is dedicated to my partner Patrick, my parents and my brother, who have provided me with absolute and unconditional support and encouragement throughout my PhD study. Furthermore, I offer my sincerest gratitude to Professor Bonggeun Kim at the Department of Economics, Seoul National University, for his unconditional support and invaluable suggestions during the difficult times. Without these very important people in my life, this thesis would not have been completed. I am also thankful for all their valuable support and dearly appreciate the friendship I obtained from Dr Duygu Yengin, Dr Prue Kerr and Dr Erkan Yalcin. In particular, Dr Duygu Yengin has given me absolute and unconditional friendship which has helped me to get through the most difficult times at Adelaide.

Abstract

During the past few decades, nonparametric models have been extensively applied to empirical studies in various fields of economics due to its flexibility for depicting any type of relationship among key economic variables. However, one of the most well-known shortfalls of the model is the *curse of dimensionality*. It can be conveniently overcome with semiparametric modelling such as partially linear (PL) models and/or single-index (SI) models. Nonetheless, the practicality of these models in the empirical studies has been hampered by the lack of appropriate estimation procedures and a method to address endogeneity. Hence the ultimate goal of this thesis is to establish a novel econometric method for estimating semiparametrics, specifically a PL model and an extended generalised partially linear single-index (EGPLSI) model, with the presence of endogeneity. Furthermore, semiparametric analysis is an important tool for analysing empirical Engel curves, which often involve endogeneity in total expenditure. We show that, our newly developed estimation procedures and methods are able to address the endogeneity problem in the semiparametric analysis of empirical Engel curves. These goals can be broken down into a few research objectives.

- (1) Firstly, this thesis aims to construct a comprehensive and systematic treatment of endogeneity in semiparametrics, given the complexity of the models containing both parametric and nonparametric components.
- (2) Secondly, it aims to develop novel estimation procedures and methods to address endogeneity in a PL model and an EGPLSI model.
- (3) Lastly, it aims to analyse the empirical demand function semiparametrically by applying the estimation procedures and methods in this thesis.

Publications arising from the thesis

- (1) Kim, N. and Saart, P. W. (2013). Estimation in Partially Linear Semiparametric Models with parametric and/or nonparametric endogeneity. *Under review at the Econometrics Journal*.
- (2) Kim, N., Saart, P. W. and Gao, J. (2013). Semi-parametric Analysis of Shape-Invariant Engel Curves with Control Function Approach. *Under review at the Journal of Econometrics*.

Contents

Acknowledgements	iii
List of Figures	x
List of Tables	xi
1 Background and Motivation	1
1.1 Introduction	1
1.2 Review of Nonparametrics and Semiparametrics in the Presence of Endogeneity	3
1.3 Review of the Empirical Engel Curves Literature	6
1.4 Research Objectives and Thesis Structure	9
2 Endogeneity in a PL Model	11
2.1 Introduction	11
2.2 Endogeneity in a PL Model	16
2.2.1 The PL Model	16
2.2.2 Endogeneity in the PL Model	18
2.2.3 Parametric Endogeneity	20
2.2.4 Nonparametric Endogeneity	23
2.3 Simulations	30
2.4 Conclusions	36
2.5 Appendix	37
2.5.1 Conditions for the PIV estimator	37
2.5.2 Conditions for Theorems 2.2.1 and 2.2.2	39
2.5.3 Proof of Theorem 2.2.1	41
2.5.4 Proof of Theorem 2.2.2	55
3 Extended Generalised Partially Linear Single-Index Model with Control Function Approach	59
3.1 Introduction	59
3.2 EGPLSI Model with/without Endogeneity	61

3.2.1	EGPLSI Model without Endogeneity	62
3.2.2	EGPLSI Model with Endogeneity	64
3.2.3	Asymptotic Properties	70
3.3	Simulation Studies	74
3.3.1	Initial Investigation	74
3.3.2	More Detailed Analysis	78
3.4	Conclusions	87
3.5	Appendix	87
3.5.1	Proofs of Theorem 3.2.1 and Corollary 3.2.2	88
3.5.2	Proof of Theorem 3.2.2	105
4	Semiparametric Analysis of Empirical Engel Curves in Australia	109
4.1	Introduction	109
4.2	The Empirical Model	111
4.2.1	A Simple Test of Endogeneity	113
4.2.2	Shape-Invariant Engel Curves	119
4.2.3	Empirical Findings	119
4.3	Conclusions	129
5	Conclusion	130
5.1	Summary	130
5.2	Future Research	132

List of Figures

3.1	$g(\cdot)$, $\iota_1(\cdot)$, $\iota_2(\cdot)$ and $\iota_3(\cdot)$	83
4.1	Kernel joint density estimates with a full bandwidth matrix.	114
4.2	Kernel estimates of conditional expectation of $\log(\text{expenditure})$ with respect to $\log(\text{income})$	115
4.3	$\log(\text{expenditure})$, m_{X_1} and η	117
4.4	Engel curves for alcohol	121
4.5	Engel curves for clothing	123
4.6	Engel curves for electricity and gas	125
4.7	Engel curves for transportation	126
4.8	Engel curves for food	127
4.9	Engel curves for other goods	128

List of Tables

2.1	The effects of endogeneity on the PL model and the appropriate estimation methods	20
2.2	Exogenous model with 2SR	31
2.3	Linear Endogeneity model with 2SR	32
2.4	Nonlinear Endogeneity model with 2SR	32
2.5	Linear Endogeneity model with 2SCF	33
2.6	Nonlinear Endogeneity model with 2SCF	33
2.7	Linear parametric endogeneity model with 2SR	34
2.8	Nonlinear parametric endogeneity model with 2SR	35
2.9	Linear parametric endogeneity model with 2SR-PIV	35
2.10	Nonlinear parametric endogeneity model with 2SR-PIV	35
3.1	GPLSI-type model with nonparametric endogeneity: Procedure 3.2.1	76
3.2	GPLSI-type model with nonparametric endogeneity: Procedure 3.2.2	76
3.3	EGPLSI-type model with nonparametric endogeneity: Procedure	
3.2.1.	77
3.4	EGPLSI-type model with nonparametric endogeneity: Procedure	
3.2.2.	77
3.5	Nonparametric exogeneity, i.e. $\iota 1$	79
3.6	Linear endogeneity, i.e. $\iota 2$	80
3.7	Nonlinear endogeneity, i.e. $\iota 3$	81
3.8	$Corr_{X_{2i}, Z_i}$	82
3.9	Linear endogeneity, i.e. $\iota 2$	85
3.10	Nonlinear endogeneity, i.e. $\iota 3$	86
4.1	Descriptive statistics.	112
4.2	Empirical results	129