

Michael Lebacher

Aus der Reihe: e-fellows.net stipendiaten-wissen

e-fellows.net (Hrsg.)

Band 2376

Estimation in Case of Endogenous Selection with Application to Wage Regression

Master's Thesis

YOUR KNOWLEDGE HAS VALUE



- We will publish your bachelor's and master's thesis, essays and papers
- Your own eBook and book - sold worldwide in all relevant shops
- Earn money with each sale

Upload your text at www.GRIN.com
and publish for free



Bibliographic information published by the German National Library:

The German National Library lists this publication in the National Bibliography; detailed bibliographic data are available on the Internet at <http://dnb.dnb.de> .

This book is copyright material and must not be copied, reproduced, transferred, distributed, leased, licensed or publicly performed or used in any way except as specifically permitted in writing by the publishers, as allowed under the terms and conditions under which it was purchased or as strictly permitted by applicable copyright law. Any unauthorized distribution or use of this text may be a direct infringement of the author s and publisher s rights and those responsible may be liable in law accordingly.

Imprint:

Copyright © 2016 GRIN Verlag
ISBN: 9783668480162

This book at GRIN:

<https://www.grin.com/document/369392>

Michael Lebacher

Aus der Reihe: e-fellows.net stipendiaten-wissen

e-fellows.net (Hrsg.)

Band 2376

Estimation in Case of Endogenous Selection with Application to Wage Regression

GRIN - Your knowledge has value

Since its foundation in 1998, GRIN has specialized in publishing academic texts by students, college teachers and other academics as e-book and printed book. The website www.grin.com is an ideal platform for presenting term papers, final papers, scientific essays, dissertations and specialist books.

Visit us on the internet:

<http://www.grin.com/>

<http://www.facebook.com/grincom>

http://www.twitter.com/grin_com

Estimation in Case of Endogenous Selection with Application to Wage Regressions

Master's Thesis

Humboldt-Universität zu Berlin
School of Business and Economics
Institute for Statistics and Econometrics
Chair of Econometrics

by

Michael Lebacher

in fulfillment of the requirements
for the degree of
Master of Science (Statistics)

Berlin, September 01, 2016

Acknowledgement

I would like to express my gratitude to Christoph Breunig for his guidance throughout this thesis. I am also grateful to the participants and instructors of the Studienabschlusssseminar for the very useful tips and remarks they gave me for this thesis. Additionally, I would like to thank Xavier D'Haultfoeuille who kindly replied to questions although being on holiday. A special gratitude to my family, my girlfriend Diana, and to my friends for their comprehension and motivation.

Abstract

This thesis addresses the problem of linear regression estimation with selectively observed response data when selection is endogenous. The approach relies critically on the existence of an instrument that is independent of the selection, conditional on potential outcomes and other covariates. A parametric two-step estimation procedure is proposed. In a first step the probability of selection is estimated employing a generalized method of moments estimator. The second step uses the estimated probability weights in order to perform an inverse probability weighted least squares estimation. Two potential estimators are presented and expressions for their asymptotic variance-covariance matrices are provided. As an extension, it is shown how the concept could be used in multiple period setup, using a pooled weighted least squares estimator. Finite sample properties are illustrated in a Monte Carlo simulation study. An empirical illustration is given, using the Survey of Health, Ageing and Retirement in Europe dataset, applying the theory to wage regressions.

Contents

List of Abbreviations	iv
List of Figures	v
List of Tables	vi
1 Introduction	1
2 Theory and Models	4
2.1 Literature Review and General Theory	4
2.2 Application to the Linear Model in a Cross-Sectional Setting	11
2.3 Testability of the Independence Condition	26
2.4 Application to the Linear Model in a Pooled OLS Setting	28
3 Simulations	36
3.1 Cross-Sectional Simulations	38
3.2 Pooled OLS Simulations	40
3.3 Sargan-Hansen J -Test Simulations	41
3.4 Conclusions from the Simulations	42
4 Data and Statistical Analysis	43
4.1 Empirical Literature Review	43
4.2 Data Description, Variables and Manipulations	45
4.3 Statistical Analysis	49
5 Conclusions	57
References	59
A Mathematical Annex	64
B Figures	74
C Tables	80

List of Abbreviations

DGP	Data generating process
FGLS	Feasible generalized least squares
GMM	Generalized method of moments
IPW	Inverse probability weighting
MAR	Missing-at-random
MCAR	Missing-completely-at-random
MNAR	Missing-not-at-random
OECD	Organisation for Economic Co-operation and Development
OLS	Ordinary least squares
SHARE	Survey of Health, Ageing and Retirement in Europe
UWLLN	Uniform weak law of large numbers

List of Figures

1	Missingness in a Two-Period Setting	30
2	Normal Distribution and Distribution of Selective Observed Variable	74
3	Distribution of Selective Observed Variable and Distribution of Weighted Se- lective Observed Variable	74
4	Simulation Density 1 (Sargan-Hansen J -Test)	75
5	Simulation Density 2 (Sargan-Hansen J -Test)	75
6	Simulation Density 3 (Sargan-Hansen J -Test)	76
7	Simulation Density 4 (Sargan-Hansen J -Test)	76
8	Moment Conditions for Wave 1	77
9	Densities of Logarithmic Wages (Before Inverse Probability Weighting)	77
10	Densities of Logarithmic Wages (After Inverse Probability Weighting)	78
11	Moment Conditions for Wave 2	78
12	Moment Conditions for Waves 1-2	79

List of Tables

1	Simulation Section: Definitions	37
2	Simulation Section: Cross-Sectional DGPs	38
3	Simulation Section: Pooled OLS DGPs	40
4	Simulation Section: Sargan-Hansen J -Test DGPs	41
5	Variables Included in the Statistical Analysis	47
6	Simulation Results 1 (Cross-Sectional)	80
7	Simulation Results 2 (Cross-Sectional)	80
8	Simulation Results 3 (Cross-Sectional)	81
9	Simulation Results 4 (Cross-Sectional)	81
10	Simulation Results 5 (Cross-Sectional)	82
11	Simulation Results 6 (Cross-Sectional)	82
12	Simulation Results (Pooled OLS)	83
13	Simulation Results (Sargan-Hansen J -Test)	83
14	Countries Included in SHARE	84
15	Summary Statistics for Wave 1	85
16	Summary Statistics for Wave 2	86
17	OLS Estimates for Available Responses in Wave 1	87
18	Probit Estimates for Response Probability in Wave 1	88
19	GMM Estimates for Response Probability in Wave 1	89
20	OLS-IPW Estimates in Wave 1	90
21	Robustness Analysis in Wave 1	91
22	OLS Estimates for Available Responses in Wave 2	92
23	Probit Estimates for Response Probability in Wave 2	93
24	GMM Estimates for Response Probability in Wave 2	94
25	OLS-IPW Estimates in Wave 2	95
26	Robustness Analysis in Wave 2	96
27	POLS Estimates for Available Responses in Waves 1-2	97
28	Probit Estimates for Response Probability in Waves 1-2	98
29	GMM Estimates for the Response Probability in Waves 1-2	99
30	OLS-IPW Estimates in Waves 1-2	100
31	Robustness Analysis in Waves 1-2	101