

multilevelPSA: An R Package for Multilevel Propensity Score Analysis

Jason M. Bryer
jbryer@bryer.org

November 11, 2011

Abstract

The use of propensity score analysis (Rosenbaum & Rubin, 1983) has gained increasing popularity for the estimation of causal effects within observational studies. However, its use in situations where data is multilevel, or clustered, is limited (Arpino & Mealli, 2008). This talk will introduce the multilevelPSA (Bryer, 2011) package for R that provides functions for estimating propensity scores for large datasets using logistic regression and conditional inference trees. Furthermore, a set of graphical functions that extends the framework of visualizing propensity score analysis introduced by Helmreich and Pruzek (2009) to multilevel analysis will be discussed. An application for estimating the effects of private schools on reading, mathematics, and science outcomes from the Programme for International Student Assessment (PISA; Organization for Economic Co-operation and Development, 2009) is provided.

Keywords: *PSA, propensity score analysis, multilevel, graphics*

1 Introduction

```
> install.packages('multilevelPSA', repos=c('http://R-Forge.R-project.org',  
+ 'http://cran.r-project.org'), dep=TRUE)  
  
> library(multilevelPSA)
```

2 Programme for International Student Achievement

```
http://www.pisa.oecd.org/  
  
> download.file('http://multilevelpsa.r-forge.r-project.org/pisa/pisa.student.Rdata',  
+ 'pisa.student.Rdata')  
> download.file('http://multilevelpsa.r-forge.r-project.org/pisa/pisa.school.Rdata',  
+ 'pisa.school.Rdata')
```

```

> load('.../.../.../pisa.student.Rdata')
> load('.../.../.../pisa.school.Rdata')
> school = school.orig[,c('COUNTRY', "CNT", "SCHOOLID",
+       "SC02Q01", #Public (1) or private (2)
+       "STRATIO" #Student-teacher ration
+ )]
> names(school) = c('COUNTRY', 'CNT', 'SCHOOLID', 'PUBPRIV', 'STRATIO')
> school$SCHOOLID = as.integer(school$SCHOOLID)

```

Table 1: Number of Private and Public Schools by Country

	Public	Private	Missing
	164	17	0
	158	4	0
	141	58	0
	217	136	0
	234	39	9
	89	189	0
	812	98	37
	174	4	0
	896	76	6
	80	103	17
	137	15	0
	97	61	0
	222	51	2
	154	4	0
	234	15	12
	231	50	4
	168	7	0
	191	12	0
	0	0	168
	202	11	13
	173	11	0
	10	140	1
	163	24	0
	112	2	17
	85	98	0
	57	87	0
	140	31	5
	987	84	26
	135	51	0
	191	8	0
	181	29	0
	99	58	0
	167	6	0

...continued from previous page

Public	Private	Missing
181	3	0
10	2	0
193	2	1
30	9	0
3	42	0
1332	200	3
50	2	0
69	113	4
153	10	0
186	4	7
136	40	12
189	51	0
166	19	0
185	29	0
88	59	6
158	1	0
212	1	0
184	3	3
167	4	0
172	17	0
336	5	0
512	359	18
159	30	0
399	22	5
204	26	0
120	32	6
44	146	0
148	17	0
169	1	0
439	17	26
154	11	0
193	39	0

Table 2: Covariates Used for Propensity Score Estimations

Variable	Name	Description
CNT	CNT	Country
SCHOOLID	SchoolId	SchoolID
StIDStd	StudentId	Student ID
ST01Q01	Grade	Grade
ST04Q01	Sex	Sex

...continued from previous page

Variable	Name	Description
ST05Q01	Attend	Attend
ST06Q01	Age	Age
ST07Q01	Repeat	Repeat
ST08Q01	Mother	At home mother
ST08Q02	Father	At home father
ST08Q03	Brother	At home brothers
ST08Q04	Sister	At home sisters
ST08Q05	GrandPa	At home grandparents
ST08Q06	Other	At home others
ST10Q01	MomEd	Mother highest schooling
ST12Q01	MomJob	Mother current job status
ST14Q01	DadEd	Father highest schooling
ST16Q01	DadJob	Father current job status
ST19Q01	Lang	Language at home
ST20Q01	Desk	Desk
ST20Q02	OwnRoom	Own room
ST20Q03	StudyPl	Study place
ST20Q04	Computer	Computer
ST20Q05	Software	Software
ST20Q06	Internet	Internet
ST20Q07	Lit	Literature
ST20Q08	Poetry	Poetry
ST20Q09	Art	Art
ST20Q10	TxtBooks	Textbooks
ST20Q12	Dict	Dictionary
ST20Q13	DishW	Dishwasher
ST20Q14	DVD	DVD
ST21Q01	CellPh	How many cellphones
ST21Q02	TVs	How many TVs
ST21Q03	nComp	How many computers
ST21Q04	nCars	How many cars
ST21Q05	nBaths	How many rooms bath or shower
ST22Q01	nBooks	How many books
ST23Q01	Reading	Reading enjoyment time
ST31Q01	EnrichLang	Enrich in test language
ST31Q02	EnrichMath	Enrich in mathematics
ST31Q03	EnrichScie	Enrich in science
ST31Q05	RemedialLang	Remedial in test language
ST31Q06	RemedialMath	Remedial in mathematics
ST31Q07	RemedialScie	Remedial in science
ST32Q01	LangLessons	Out of school lessons in test language
ST32Q02	MathLessons	Out of school lessons maths
ST32Q03	ScieLessons	Out of school lessons in science

...continued from previous page

Variable	Name	Description
----------	------	-------------