`multilevelPSA`: An R Package for Estimating and Visualizing Multilevel Propensity Score Models

Jason Bryer

University at Albany

November 14, 2011

Agenda

Installing multilevelPSA

The `multilevelPSA` package is currently under development and is available on github. The `devtools` package provides a function to install R packages directly from github.

```
> library(devtools)
> install_github('multilevelPSA', 'jbryer')
```

Installing multilevelPSA

The `multilevelPSA` package is currently under development and is available on github. The `devtools` package provides a function to install R packages directly from github.

```
> library(devtools)
> install_github('multilevelPSA', 'jbryer')
```

Once installed from github, it can be loaded just like any other R package.

```
> library(multilevelPSA)
> ls('package:multilevelPSA')
```

```
 [1] "GeomRugAlt"                    "geom_rug_alt"
 [3] "getPropensityScores"           "getStrata"
 [5] "missingPlot"                   "multilevelCtree"
 [7] "multilevelLR"                  "multilevelPSA"
 [9] "plot.multilevel.distribution" "plotcirc.multilevel.psa"
[11] "plotpsa.multilevel.psa"        "treeHeat"
```

PISA: Programme for International Student Assessment

The Economic Co-operation and Development (OECD) began assessing student achievement in 2003 with the Programme of International Student Assessment (PISA; `http://www.pisa.oecd.org/`). In 2009 they evaluated students at the end of secondary school (or equivalent depending on country) in mathematics, reading, and science. Data is freely available on their website but an R data is made available with the `multilevelPSA` package. There are two data frames, `student.orig` and `school.orig` and are loaded using the `data` command.

```
> data(pisa.student)
> #names(student.orig)
> nrow(student.orig)
[1] 475460
> ncol(student.orig)
[1] 305
> data(pisa.school)
> #names(school.orig)
> nrow(school.orig)
[1] 17145
> ncol(school.orig)
[1] 247
```

Covariates

Variable	ShortDesc	Desc
CNT	CNT	Country
SCHOOLID	SchoolId	SchoolID
StIDStd	StudentId	Student ID
ST01Q01	Grade	Grade
ST04Q01	Sex	Sex
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

Table: Covariates Used for Propensity Score Estimations

Covariates (cont.)

Variable	ShortDesc	Desc
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

Table: Covariates Used for Propensity Score Estimations

Covariates (cont.)

Variable	ShortDesc	Desc
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

Table: Covariates Used for Propensity Score Estimations

Setup School Data

```
> 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)
```

Number of Private and Public Schools by Country

■results=tex■ t = table(school$COUNTRY$, $school$PUBPRIV, useNA='ifany') x = xtable(t[1:22,], caption='Number of Private and Public Schools by Country', label='ppxtab') print(x, include.rownames=FALSE, include.colnames=TRUE)

Thank You

Jason Bryer (jason@bryer.org)

`https://github.com/jbryer/multilevelPSA`