MRC Integrative Epidemiology Unit

University of BRISTOL

R and Stata packages for one-sample Mendelian randomization analyses: OneSampleMR and ivonesamplemr

Tom Palmer tom.palmer@bristol.ac.uk

20th January 2022

Outline

Overview

OneSampleMR R package

ivonesamplemr Stata package

Common pitfalls

Discussion

Overview

- Motivation
  - Code from appendix of Clarke et al. (2015) was in R scripts and Stata do-files
  - Over the summer some interesting in using it
- One-sample (a.k.a. individual level) Mendelian randomization (MR) data
  - i.e., not genotype summary level (a.k.a. two-sample) data!
- Aim
  Collection of useful functions and instrumental variable (IV) estimators (that aren't available elsewhere)

What's available elsewhere?

- Stata
  - Official Stata IV commands begin `iv`*something* e.g., `ivregress`, `ivprobit`, `ivpoisson`, `gmm`
  - User-written: `ivreg2`, `ivpois`, and many more ...
- R
  - **sem**
  - **AER**/**ivreg**
  - **ivtools** – excellent but only allows a single instrument (Sjolander et al. 2020)
  - **nlmr** (Staley et al. 2017)
  - and many more ...
- IEU software website https://mrcieu.github.io/
- Chris Moreno-Stokoe webpage https://www.morenostok.io/mrsoftwarelist.html

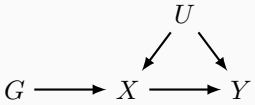
What's missing?

- Some binary outcome IV estimators
  - Structural Mean Models (SMMs); Additive SMM, Multiplicative SMM, Logistic SMM
  - Two-stage predictor substitution (TSPS)
  - Two-stage residual inclusion (TSRI)
  - Some nonlinear estimators, e.g., Burgess, Davies, et al. (2014)

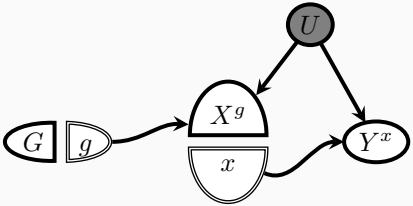
Summary of observational and IV estimators

Outcome	Link fn	Observational estimator	IV estimator
Continuous	Identity	Linear regression	Linear IV: TSLS, ...
Binary	Identity	Binomial regression with identity link	Additive SMM
Binary/Cat.	Log (add.)	Log-binomial/Poisson regresssion	TSPS, TSRI
Binary/Cat.	Log (mult.)	Gamma regression with log link	Multiplicative SMM, TSPS, TSRI
Binary	Logit	Logistic regression	Logistic SMM, TSPS, TSRI

- View potential outcomes on a/*the* IV Directed Acyclic Graph (DAG)



- using a Single World Intervention Graph (SWIG) (Swanson et al. 2018)



Features common to both OneSampleMR and ivonesamplemr packages

- Multiplicative SMM (Hernán et al. 2006)
- TSPS and TSRI estimators with a choice of second stage link functions (Terza et al. 2008)
- The functions use the generalised method of moments (GMM) estimation approach described in Clarke et al. (2015)

OneSampleMR R package

- Website https://remlapmot.github.io/OneSampleMR/
- Install from CRAN

```
install.packages("OneSampleMR")
```

OneSampleMR R package

cran.r-project.org/web/packages/OneSampleMR/index.html

OneSampleMR: One Sample Mendelian Randomization and Instrumental Variable Analyses

Useful functions for one-sample (individual level data) Mendelian randomization and instrumental variable analyses. The package includes implementations of: the Sanderson and Windmeijer (2016) <doi:10.1016/j.jeconom.2015.06.004> conditional F-statistic, the multiplicative structural mean model Hernán and Robins (2006) <doi:10.1097/01.ede.0000222409.00878.37>, and two-stage predictor substitution and two-stage residual inclusion estimators explained by Terza et al. (2008) <doi:10.1016/j.jhealeco.2007.09.009>.

Version:	0.1.0
Depends:	R (≥ 3.6.0)
Imports:	Formula, gtools, ivpack, lmtest, mam
Suggests:	AER, harvey, rvtools, knitr, life, rmarkdown, testthat (≥ 3.0.0)
Published:	2021-11-12
Author:	Tom Palmer [aut, cre], Wes Spiller [aut], Eleanor Sanderson [aut]
Maintainer:	Tom Palmer <remlapmot at hotmail.com>
BugReports:	https://github.com/remlapmot/OneSampleMR/issues/
License:	GPL (≥ 3)
URL:	https://github.com/remlapmot/OneSampleMR
NeedsCompilation:	no
Materials:	README NEWS
CRAN checks:	OneSampleMR results

Documentation:

Reference manual:	OneSampleMR.pdf
Vignettes:	Comparison fits of the multiplicative structural mean model Comparison of conditional F-statistics

Downloads:

Package source:	OneSampleMR_0.1.0.tar.gz
Windows binaries:	r-devel: OneSampleMR_0.1.0.zip, r-release: OneSampleMR_0.1.0.zip, r-oldrel: OneSampleMR_0.1.0.zip
macOS binaries:	r-release (arm64): OneSampleMR_0.1.0.tgz, r-release (x86_64): OneSampleMR_0.1.0.tgz, r-oldrel: OneSampleMR_0.1.0.tgz

Linking:

Please use the canonical form https://CRAN.R-project.org/package=OneSampleMR to link to this page.

OneSampleMR R package

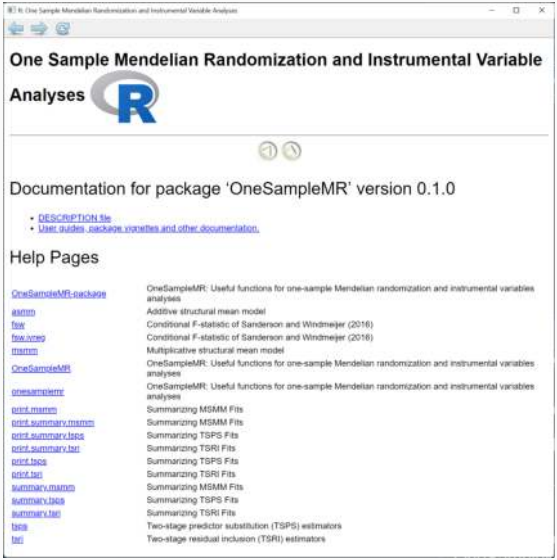
- For IV estimators syntax follows syntax of AER `::ivreg()`/`ivreg::ivreg()`

```
msmm(outcome ~ exposure + confndrs | instruments + confndrs, ...)
```

- Sanderson et al. (2016) conditional F-statistic (already in Stata's user-written `ivreg2` and **lfe** R package); `fsw()`
- Package helpfile

```
help(package = "OneSampleMR")
```

OneSampleMR R package



R: One Sample Mendelian Randomization and Instrumental Variable Analyses

One Sample Mendelian Randomization and Instrumental Variable Analyses

Documentation for package 'OneSampleMR' version 0.1.0

- DESCRIPTION file
- User guides, package vignettes and other documentation.

Help Pages

OneSampleMR-package	OneSampleMR: Useful functions for one-sample Mendelian randomization and instrumental variables analyses
asmm	Additive structural mean model
fsw	Conditional F-statistic of Sanderson and Windmeijer (2016)
fsw.ivreg	Conditional F-statistic of Sanderson and Windmeijer (2016)
msmm	Multiplicative structural mean model
OneSampleMR	OneSampleMR: Useful functions for one-sample Mendelian randomization and instrumental variables analyses
onesamplemr	OneSampleMR: Useful functions for one-sample Mendelian randomization and instrumental variables analyses
print.msmm	Summarizing MSMM Fits
print.summary.msmm	Summarizing MSMM Fits
print.summary.tsps	Summarizing TSPS Fits
print.summary.tsri	Summarizing TSRI Fits
print.tsps	Summarizing TSPS Fits
print.tsri	Summarizing TSRI Fits
summary.msmm	Summarizing MSMM Fits
summary.tsps	Summarizing TSPS Fits
summary.tsri	Summarizing TSRI Fits
tsps	Two-stage predictor substitution (TSPS) estimators
tsri	Two-stage residual inclusion (TSRI) estimators

ivonesamplemr Stata package

- Repository https://github.com/remlapmot/ivonesamplemr
- Install with

```
net install github, from("https://haghish.github.io/github/")
github install remlapmot/ivonesamplemr
```

- Command syntax follows `ivregress` syntax

```
ivmsmm outcome confounders (exposure = instruments), options
```

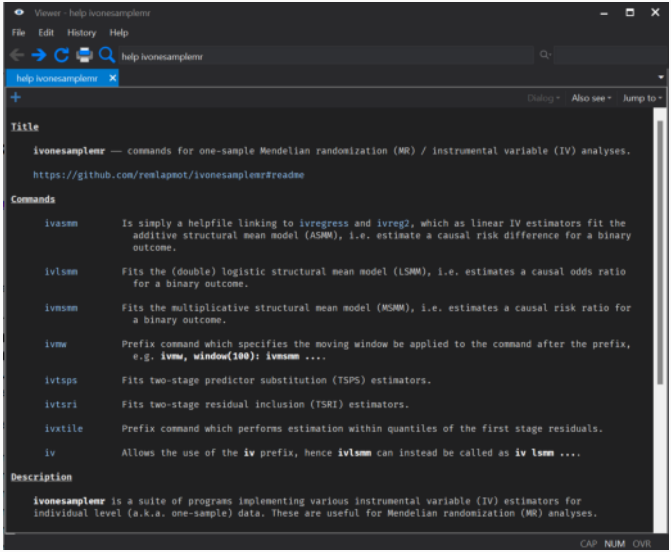
- All commands are named `ivsomething`
- Logistic SMM; `ivlsmm`

ivonesamplemr Stata package

- Moving (sliding/rolling) window of rank ordered first stage residuals (Burgess, Davies, et al. 2014); `ivmw:` prefix command
- Quantiles of first stage residuals (Burgess, Davies, et al. 2014); `ivxtile:` prefix command
- Package helpfile

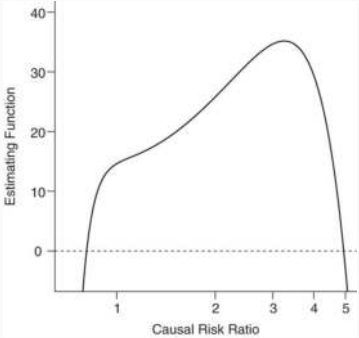
```
help ivonesamplemr
```

ivonesamplemr Stata package



Common pitfalls

- Local minima – even though GMM reports model has converged (Burgess, Granell, et al. 2014)



Common pitfalls

- Extreme estimates, e.g., estimated causal odds ratio of 10
- Very large sample sizes – hanging/non-convergence (write a Julia package??)

Discussion

- R and Stata packages for one-sample (a.k.a. individual level data) MR analyses
- Binary outcome IV estimators (SMMs/TSPS/TSRI)
- Some nonlinear estimators
- Example of use: Madley-Dowd et al. preprint
- Alternative: split sample and use two-sample methods (Burgess et al. 2016)
- **OneSampleMR** was included in the R Views November 2021 “Top 40” New CRAN Packages
- TODO - what would be useful to add?
    - **nlmr** in Stata
    - Option in TSPS/TSRI fns to automate estimation in case-control studies (i.e., first stage fitted only in controls)

References

Burgess, S., N. M. Davies, and S. G. Thompson. 2014. “Instrumental variable analysis with a nonlinear exposure–outcome relationship.” *Epidemiology* 25 (6): 877–885. https://doi.org/10.1097/EDE.0000000000000161.

———. 2016. “Bias due to participant overlap in two-sample Mendelian randomization.” *Genetic Epidemiology* 40 (7): 597–608. https://doi.org/10.1002/gepi.21998.

Burgess, S., R. Granell, T. M. Palmer, J. A. C. Sterne, and V. Didelez. 2014. “Lack of Identification in Semiparametric Instrumental Variable Models With Binary Outcomes.” *American Journal of Epidemiology* 180 (1): 111–119. https://doi.org/10.1093/aje/kwu107.

Clarke, P. S., T. M. Palmer, and F. Windmeijer. 2015. “Estimating Structural Mean Models with Multiple Instrumental Variables Using the Generalised Method of Moments.” *Statistical Science* 30 (1): 96–117. https://doi.org/10.1214/14-STS503.

Hernán, M. A., and J. M. Robins. 2006. “Instruments for causal inference: an Epidemiologist’s dream?” *Epidemiology* 17 (4): 360–372. https://doi.org/10.1097/01.ede.0000222409.00878.37.

References

Sanderson, E., and F. Windmeijer. 2016. “A weak instrument F-test in linear IV models with multiple endogenous variables.” *Journal of Econometrics* 190 (2): 212–221. https://doi.org/10.1016/j.jeconom.2015.06.004.

Sjolander, A., E. Dahlqwist, and T. Martinussen. 2020. *ivtools: Instrumental Variables.* R package version 2.3.0.

Staley, J., and S. Burgess. 2017. “Semiparametric methods for estimation of a non-linear exposure-outcome relationship using instrumental variables with application to Mendelian randomization.” *Genetic Epidemiology* 41 (4): 341–352. https://doi.org/10.1002/gepi.22041.

Swanson, S. A., M. A. Hernán, M. Miller, J. M. Robins, and T. S. Richardson. 2018. “Partial Identification of the Average Treatment Effect Using Instrumental Variables: Review of Methods for Binary Instruments, Treatments, and Outcomes.” PMID: 31537952, *Journal of the American Statistical Association* 113 (522): 933–947. https://doi.org/10.1080/01621459.2018.1434530.

Terza, J. V., A. Basu, and P. J. Rathouz. 2008. “Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling.” *Journal of Health Economics* 27 (3): 531–543. https://doi.org/10.1016/j.jhealeco.2007.09.009.