yotover: An R package to support analysis and reproduce results from An Advanced Guide to Trade Policy Analysis

30 Oct 2020

Summary

The tools created by the open source community have greatly eased the burden of documenting work in econometrics (Koenker and Zeileis 2009). Following this statement, and the replication case studies shown in Koenker and Zeileis (2009), we decided to reproduce the results from Yotov et al. (2016) by using the R programming language. These results, which we could fully replicate because the authors provided both the data and documented code, were originally obtained in Stata, leaving space for contributions such as this to contribute to an effective communication of research throughout distribution of open source software and open formats data. Previous work in the R community, such as Wölwer, Breßlein, and Burgard (2018), Wölwer et al. (2018) and Porto (2020) have already started easing estimation methods for gravity models in R.

`yotover` is an R package designed to support analysis of the datasets accompanying the book An Advanced Guide to Trade Policy Analysis (Yotov et al. 2016). To facilitate further analysis of the original datasets, the `yotover` package creates a local, on-disk embedded database (Raasveldt and Mühleisen 2019) with the corresponding data. This avoids the need for users to pre-process the data or load the multi-gigabyte dataset into memory. The DuckDB back-end allows high-performance querying and is accessible via a `DBI`- and `dplyr`-compatible interface familiar to most R users (R Special Interest Group on Databases (R-SIG-DB), Wickham, and Müller 2018; Wickham et al. 2019). For users of the RStudio integrated development environment (RStudio Team 2015), the package also provides an interactive pane for exploring the database and previewing data.

Acknowledgements

The authors were funded by the support of the United Nations Economic and Social Commission for Asia and the Pacific.

References

Koenker, Roger, and Achim Zeileis. 2009. "On Reproducible Econometric Research." *Journal of Applied Econometrics* 24 (5): 833–47.

Porto, Massimiliano. 2020. *Using R for Trade Policy Analysis: R Codes for the Unctad and Wto Practical Guide*. Springer Nature.

Raasveldt, Mark, and Hannes Mühleisen. 2019. "DuckDB: An Embeddable Analytical Database." In *Proceedings of the 2019 International Conference on Management of Data*, 1981–4.

R Special Interest Group on Databases (R-SIG-DB), Hadley Wickham, and Kirill Müller. 2018. *DBI: R Database Interface*. https://CRAN.R-project.org/package=DBI.

RStudio Team. 2015. *RStudio: Integrated Development Environment for R*. Boston, MA, USA: RStudio, Inc. http://www.rstudio.com/.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2019. *dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wölwer, Anna-Lena, Martin Breßlein, and Jan Pablo Burgard. 2018. "Gravity Models in R." *Austrian Journal of Statistics* 47 (4): 16–35. https://doi.org/10.17713/ajs.v47i4.688.

Wölwer, Anna-Lena, Jan Pablo Burgard, Joshua Kunst, and Mauricio Vargas. 2018. "Gravity: Estimation Methods for Gravity Models in R." *Journal of Open Source Software* 3 (31): 1038.

Yotov, Yoto V, Roberta Piermartini, José-Antonio Monteiro, and Mario Larch. 2016. *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. World Trade Organization Geneva.