

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

Alexey Kravchenko¹ and Mauricio Vargas Sepúlveda²

¹ United Nations Economic and Social Commission for Asia and the Pacific ² Pontifical Catholic University of Chile

DOI:

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted:

Published:

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

Summary

The tools created by the open source community have greatly eased the burden of documenting work in econometrics according to Koenker & Zeileis (2009). We decided to try to reproduce the results from Yotov, Piermartini, Monteiro, & Larch (2016) by using the R programming language to see if we could reproduce its results as it was done with other studies in the aforementioned article.

`yotover` is an R package designed to organize our replication work, which support analysis of the datasets accompanying the book *An Advanced Guide to Trade Policy Analysis* (Yotov et al., 2016). We followed the work of Ross, Eskew, & Ray (2019) in ecology, and we were able to perform a full replication of the results, in part because the authors provided linked analytical data and documented executable Stata code, leaving space for contributions such as this in particular. How Yotov et al. (2016) provided their results contributes to an effective communication of research, and in the reproducibility spectrum their work is closer to the full replication gold standard described in Peng (2011) than to publication only.

To facilitate further analysis of the original datasets, the `yotover` package creates a local, on-disk embedded database with the corresponding data. This avoids the need for users to pre-process the data, load the multi-gigabyte dataset into memory, and it respects CRAN packages policies regarding datasets. The DuckDB back-end (Raasveldt & Mühleisen, 2019) 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, & Müller, 2018; Wickham, François, Henry, & Müller, 2019).

This works aims at making the reproduction of the original results more accessible throughout the distribution of both open source software and open formats data. Previous work in the R community, such as Wölwer, Breßlein, & Burgard (2018), Wölwer, Burgard, Kunst, & Vargas (2018) and Porto (2020) have already started easing estimation methods for gravity models in R.

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, considering that this software in particular aims at economists mostly, who are often accustomed to Excel and Stata instead of R, Python and other tools such as SQL.

Acknowledgements

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

References

- Koenker, R., & Zeileis, A. (2009). On reproducible econometric research. *Journal of Applied Econometrics*, 24(5), 833–847.
- Peng, R. D. (2011). Reproducible research in computational science. *Science*, 334(6060), 1226–1227.
- Porto, M. (2020). *Using r for trade policy analysis: R codes for the unctad and wto practical guide*. Springer Nature.
- Raasveldt, M., & Mühleisen, H. (2019). DuckDB: An embeddable analytical database. In *Proceedings of the 2019 international conference on management of data* (pp. 1981–1984).
- Ross, N., Eskew, E. A., & Ray, N. (2019). Citesdb: An r package to support analysis of cites trade database shipment-level data. *Journal of Open Source Software*, 4(37), 1483.
- R Special Interest Group on Databases (R-SIG-DB), Wickham, H., & Müller, K. (2018). *DBI: R database interface*. Retrieved from <https://CRAN.R-project.org/package=DBI>
- RStudio Team. (2015). *RStudio: Integrated development environment for R*. Boston, MA, USA: RStudio, Inc. Retrieved from <http://www.rstudio.com/>
- Wickham, H., François, R., Henry, L., & Müller, K. (2019). *dplyr: A grammar of data manipulation*. Retrieved from <https://CRAN.R-project.org/package=dplyr>
- Wölwer, A.-L., Breßlein, M., & Burgard, J. P. (2018). Gravity models in r. *Austrian Journal of Statistics*, 47(4), 16–35. doi:10.17713/ajs.v47i4.688
- Wölwer, A.-L., Burgard, J. P., Kunst, J., & Vargas, M. (2018). Gravity: Estimation methods for gravity models in r. *Journal of Open Source Software*, 3(31), 1038.
- Yotov, Y. V., Piermartini, R., Monteiro, J.-A., & Larch, M. (2016). *An advanced guide to trade policy analysis: The structural gravity model*. World Trade Organization Geneva.