The Journal of Open Source Software

ARDL: An R package for the analysis of level relationships

Kleanthis Natsiopoulos[1] **and Nickolaos G. Tzeremes**[1]

1 Department of Economics, University of Thessaly, 28th October, 78, 38333, Volos, Greece

DOI: 10.21105/joss.03496

Software

- Review
- Repository
- Archive

Editor: Sebastian Benthall

Reviewers:

- @ha0ye
- @jacobsoj

Submitted: 03 February 2021
Published: 15 July 2021

Summary

Autoregressive Distributed Lag (ARDL) and Error Correction Models (ECM) are widely used in various economic applications as they are very flexible. Also, these models are used in the context of cointegration analysis as a platform to test and analyze the levels (long-run) relationship between variables. One of the most popular such tests is the bounds test proposed by Pesaran et al. (2001) which allows testing for cointegration while at the same time estimates the level relationship.

Statement of need

`ARDL` (Natsiopoulos & Tzeremes, 2021) is an `R` package that aims to help users in the modeling process of ARDL and ECM and it also provides the tools towards the bounds test for cointegration. `ARDL` is implemented in such a way that researchers can use it as a full featured tool for this specific type of analysis and students of all levels can be aware of how each piece of code works through the analytical manual and the examples which cover every functionality of the package. A recent example is Qiu et al. (2021) where the software was used to forecast tourist arrivals through an ARDL model, amonth other methods.

State of the field

`ARDL` distinguishes itself from other related `R` packages like `dLagM` (Demirhan, 2020) and `dynamac` (Jordan & Philips, 2020) in the sense that it is specifically designed to address a particular problem throughout every phase of modeling, testing and interpretation. It includes a rich set of dedicated tools, accompanied by an analytical manual that describes the mathematical process of every function.

On the other hand, `dLagM` and `dynamac` provide additional plotting functionalities and post-estimation diagnostics (serial correlation tests etc). The design of the `ARDL` packages differs from the pre-mentioned packages as it is not an *all-in-one* package. It is designed so that every exported object is of commonly used `R` classes so that it can be easily combined with other packages, each of which are also dedicated to a specific part of the post-estimation part of the analysis.

In addition, the `ARDL` packages natively supports time-series data, thus sub-sample and balanced sample estimations are possible. It also allows to specify any level of significance for the bounds F-test and t-tests, and includes p-values and exact sample critical values for any possible combination. Also, the estimation of long-run, short-run and interim multipliers accompanied by standard errors and p-values as well as the cointegrating equation are some of the available features.

The Journal of Open Source Software

Breaking down the ARDL package

Model estimation

The package does not explicitly connects the modeling with the bounds test, as the ARDL and ECM model may well be used independently in other research approaches too. This way each function is dedicated to perform a very specific part of the whole process.

An $ARDL(p, q_1, ..., q_k)$ model can be fully specified by its order. The `ardl` function implements the following formula:

$$y_t = c_0 + c_1 t + \sum_{i=1}^{p} b_{y,i} y_{t-i} + \sum_{j=1}^{k} \sum_{l=0}^{q_j} b_{j,l} x_{j,t-l} + \epsilon_t \tag{1}$$

An Unrestricted ECM (UECM) has a 1:1 relationship with the $ARDL(p, q_1, ..., q_k)$ model. The `uecm` function estimates it using the formula:

$$\Delta y_t = c_0 + c_1 t + \pi_y y_{t-1} + \sum_{j=1}^{k} \pi_j x_{j,t-1} + \sum_{i=1}^{p-1} \psi_{y,i} \Delta y_{t-i} + \sum_{j=1}^{k} \sum_{l=1}^{q_j-1} \psi_{j,l} \Delta x_{j,t-l} + \sum_{j=1}^{k} \omega_j \Delta x_{j,t} + \epsilon_t \tag{2}$$

The Restricted ECM (RECM) can also be fully described by the order of the underlying ARDL and the case for potential restriction in the deterministic parameters. The `recm` function performs the modeling using:

$$\Delta y_t = c_0 + c_1 t + \sum_{i=1}^{p-1} \psi_{y,i} \Delta y_{t-i} + \sum_{j=1}^{k} \sum_{l=1}^{q_j-1} \psi_{j,l} \Delta x_{j,t-l} + \sum_{j=1}^{k} \omega_j \Delta x_{j,t} + \pi_y ECT_t + \epsilon_t \tag{3}$$

Under Case 1:

$$\begin{aligned} & c_0 = c_1 = 0 \\ & ECT = y_{t-1} - (\sum_{j=1}^{k} \theta_j x_{j,t-1}) \end{aligned} \tag{4}$$

Under Case 2:

$$\begin{aligned} & c_0 = c_1 = 0 \\ & ECT = y_{t-1} - (\mu + \sum_{j=1}^{k} \theta_j x_{j,t-1}) \end{aligned} \tag{5}$$

Under Case 3:

$$
\begin{aligned}
& c_1 = 0 \\
& ECT = y_{t-1} - (\sum_{j=1}^{k} \theta_j x_{j,t-1})
\end{aligned} \tag{6}
$$

Under Case 4:

$$
\begin{aligned}
& c_1 = 0 \\
& ECT = y_{t-1} - (\delta(t-1) + \sum_{j=1}^{k} \theta_j x_{j,t-1})
\end{aligned} \tag{7}
$$

Under Case 5:

$$
ECT = y_{t-1} - (\sum_{j=1}^{k} \theta_j x_{j,t-1}) \tag{8}
$$

Model relationships

Figure 1 shows the relationships between the regression model objects. The bounds tests and all the other functions of the package also have this kind of interconnectivity as they inherit all the necessary information.

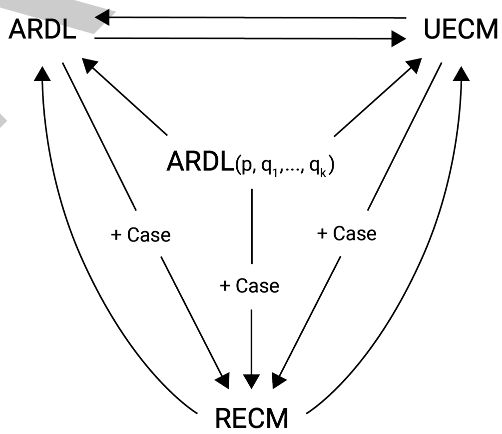

Figure 1: Interconnection between ARDL, UECM and RECM.

Where $ARDL(p, q_1, ..., q_k)$ represent the ARDL order and Case is the restriction of the deterministic parameters. When an arrow points from one model or information (order or case) to another model, it can be interpreted as the first one can fully describe the second one. When it crosses the case it means that this information is also needed.

Bounds test

The functions `bounds_f_test` and `bounds_t_test` return a typical `htest` object and they perform a Wald or t test respectively on an UECM. The input of the functions can also be an ARDL model as they are interconnected as described above. The hypothesis tests for the bounds F-test and t-test are based on Equation 2. For the bounds F-test:

Cases 1, 3, 5:

$$\begin{aligned}\mathbf{H_0} :& \pi_y = \pi_1 = \cdots = \pi_k = 0 \\ \mathbf{H_1} :& \pi_y \neq \pi_1 \neq \cdots \neq \pi_k \neq 0\end{aligned} \tag{9}$$

Cases 2:

$$\begin{aligned}\mathbf{H_0} :& \pi_y = \pi_1 = \cdots = \pi_k = c_0 = 0 \\ \mathbf{H_1} :& \pi_y \neq \pi_1 \neq \cdots \neq \pi_k \neq c_0 \neq 0\end{aligned} \tag{10}$$

Cases 4:

$$\begin{aligned}\mathbf{H_0} :& \pi_y = \pi_1 = \cdots = \pi_k = c_1 = 0 \\ \mathbf{H_1} :& \pi_y \neq \pi_1 \neq \cdots \neq \pi_k \neq c_1 \neq 0\end{aligned} \tag{11}$$

Finally, the bounds t-test can be represented as:

$$\begin{aligned}\mathbf{H_0} :& \pi_y = 0 \\ \mathbf{H_1} :& \pi_y \neq 0\end{aligned} \tag{12}$$

Making inference after cointegration

After the modeling part and if a cointegrating relationship can be established, the long-run (but also the short-run or interim) multipliers can be computed using the `multipliers` function. Note that it is irrelevant whether they are estimated based on an ARDL (Equation 1) or an UECM (Equation 2), in terms of the results. The formulas for the multipliers for the constant and the linear trend are:

When the input is an ARDL:

$$\mu = \frac{c_0}{1 - \sum_{i=1}^{p} b_{y,i}} \tag{13}$$

$$\delta = \frac{c_1}{1 - \sum_{i=1}^{p} b_{y,i}} \tag{14}$$

When the input in an UECM:

$$\mu = \frac{c_0}{-\pi_y} \tag{15}$$

$$\delta = \frac{c_1}{-\pi_y} \tag{16}$$

The relationships for the short-run multipliers are:

When the input is an ARDL:

$$\frac{\partial y_t}{\partial x_{j,t}} = \frac{b_{j,0}}{1 - \sum_{i=1}^{p} b_{y,i}} \qquad \forall j = 1, \dots, k \tag{17}$$

When the input is an UECM:

$$\frac{\partial y_t}{\partial x_{j,t}} = \frac{\omega_j}{-\pi_y} \qquad \forall j = 1, \dots, k \tag{18}$$

The relationships for the interim multipliers are:

When the input is an ARDL:

$$\frac{\partial y_{t+s}}{\partial x_{j,t}} = \frac{\sum_{l=1}^{s} b_{j,l}}{1 - \sum_{i=1}^{p} b_{y,i}} \qquad \forall j = 1, \dots, k \qquad s \in \{0, \dots, q_j\} \tag{19}$$

When the input is an UECM:

$$\frac{\partial y_{t+s}}{\partial x_{j,t}} = \frac{\pi_j + \psi_{j,s}}{-\pi_y} \qquad \forall j = 1, \dots, k \qquad s \in \{1, \dots, q_j - 1\} \tag{20}$$

The relationships for the long-run multipliers are:

When the input is an ARDL:

$$\frac{\partial y_{t+\infty}}{\partial x_{j,t}} = \theta_j = \frac{\sum_{l=0}^{q_j} b_{j,l}}{1 - \sum_{i=1}^{p} b_{y,i}} \qquad \forall j = 1, \dots, k \tag{21}$$

When the input is an UECM:

$$\frac{\partial y_{t+\infty}}{\partial x_{j,t}} = \theta_j = \frac{\pi_j}{-\pi_y} \qquad \forall j = 1, \dots, k \tag{22}$$

Lastly, the cointegrating relationship vector can be constructed using the function `coint_eq`. The formula is:

Cases 1, 3, 5:

$$CointEq_t = \sum_{j=1}^{k} \theta_j x_{j,t} \quad \forall j = 1, \dots, k \tag{23}$$

Case 2:

$$CointEq_t = \mu + \sum_{j=1}^{k} \theta_j x_{j,t} \quad \forall j = 1, \dots, k \tag{24}$$

Case 4:

$$CointEq_t = \delta + \sum_{j=1}^{k} \theta_j x_{j,t} \quad \forall j = 1, \dots, k \tag{25}$$

Conclusion

Using the intuitive API of the `ARDL` package, even the most complex ARDL, UECM or RECM model can be fully specified using the $ARDL(p, q_1, ..., q_k)$ order and the case for the restriction of the constant and linear deterministic parameters. The interconnection between the objects allows for further post-estimation testing. Finally, the detailed estimation formulas are provided in order to enhance the understanding of every step of the analysis giving a reference point for the equations in publications.

Acknowledgments

We want to acknowledge Nikolaos Chatsios for designing the package's logo and the diagram in Figure 1.

References

Demirhan, H. (2020). dLagM: An R package for distributed lag models and ARDL bounds testing. *PLoS ONE*, *15*(2), e0228812. https://doi.org/10.1371/journal.pone.0228812

Jordan, S., & Philips, A. Q. (2020). *Dynamac: Dynamic simulation and testing for single-equation ARDL models*. https://CRAN.R-project.org/package=dynamac

Natsiopoulos, K., & Tzeremes, N. (2021). *ARDL: ARDL, ECM and bounds-test for cointegration*. https://CRAN.R-project.org/package=ARDL

Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, *16*(3), 289–326. https://doi.org/10.1002/jae.616

Qiu, R. T. R., Wu, D. C., Dropsy, V., Petit, S., Pratt, S., & Ohe, Y. (2021). TOURIST ARRIVAL FORECAST AMID COVID-19: A perspective from the Asia and Pacific team. *Annals of Tourism Research*, 103155. https://doi.org/https://doi.org/10.1016/j.annals.2021.103155