ardl: Estimating autoregressive distributed lag and equilibrium correction models

Sebastian Kripfganz
University of Exeter Business School
Exeter, United Kingdom
S.Kripfganz@exeter.ac.uk

Daniel C. Schneider
Max Planck Inst. for Demographic Research
Rostock, Germany
schneider@demogr.mpg.de

Abstract. We present a Stata package for the estimation of autoregressive distributed lag (ARDL) models in a time-series context. The `ardl` command can be used to estimate an ARDL model with the optimal number of autoregressive and distributed lags based on the Akaike or Schwarz/Bayesian information criterion. The regression results can be displayed in the ARDL levels form or in the error-correction representation of the model. The latter separates long-run and short-run effects and is available in two different parameterizations of the long-run (cointegrating) relationship. The popular bounds testing procedure for the existence of a long-run levels relationship is implemented as a postestimation feature. Comprehensive critical values and approximate p-values obtained from response-surface regressions facilitate statistical inference.

Keywords: ardl, autoregressive distributed lag model, error-correction model, bounds test, long-run relationship, cointegration, time-series data

1 Introduction

Real-world phenomena are often characterized by complex relationships. Some observed variables might exhibit erratic behavior in the short run, but tend to co-move in a stable and predictable way with other variables over longer time horizons. Attempting to empirically uncover such long-run/equilibrium relationships is tantamount to separating them from the overlaid short-run dynamics. This separation then allows finding evidence for or against an equilibrium relationship, which is often at the heart of a research question. It also allows to analyze the short-term fluctuations around the equilibrium, which can be valuable in its own right, for example, when conducting forecasting exercises or dynamic simulations.

When we observe the variables of interest over a sufficiently long stretch of consecutive time periods, multi-equation vector autoregressive (VAR) models are commonly used to assess their dynamic relationships. When we have reasons to assume that there is a natural ordering of the variables, such that there is no contemporaneous feedback from a response variable to the other variables in the system, a single-equation autoregressive distributed lag (ARDL) model can simplify the analysis and facilitate more efficient inference.[1]

ARDL models have a wide range of possible applications. They are extensively

1. Occasionally, the abbreviation ADL is used in the literature instead of ARDL.

used in studies analyzing the linkages of polution and energy consumption to economic growth (Fatai et al. 2004; Narayan and Smyth 2005; Wolde-Rufael 2005, 2006; Ang 2007; Halicioglu 2009; Jalil and Mahmud 2009; Ozturk 2010; Payne 2010; Zhang et al. 2015; Ntanos et al. 2018; Bekun et al. 2019, and many more). Relationships with economic growth have also been investigated for foreign direct investment and trade (Oteng-Abayie and Frimpong 2006; Belloumi 2014), infrastructure (Fedderke et al. 2006), immigration (Morley 2006), tourism (Katircioglu 2009; Wang 2009; Song et al. 2011), and stock market development (Enisan and Olufisayo 2009).

Other examples include the nexus between wages, productivity, and unemployment (Pesaran et al. 2001), savings and investment (Narayan 2005), exchange rates and trade (Bahmani-Oskooee and Brooks 1999; De Vita and Abbott 2004), exchange rates and monetary policy (Frankel et al. 2004; Shambaugh 2004; Obstfeld et al. 2005), financial development and inequality (Ang 2010), bank lending and property prices (Davis and Zhu 2011), financial reforms and credit growth (Adeleye et al. 2018), and the interdependencies among stock price indices and commodity prices (Narayan et al. 2004; Sari et al. 2010; Büyükşahin and Robe 2014), as well as cryptocurrencies (Ciaian et al. 2016, 2018), to list only a few.

The ARDL model can be conveniently reparameterized in so-called error-correction (EC) form, which disentangles the long-run relationship from the short-run dynamics. When the variables are nonstationary – to be precise: integrated of order 1 – the long-run relationship embedded in an EC model corresponds to a cointegrating relationship (Engle and Granger 1987; Hassler and Wolters 2006). Testing for cointegration in such a setup therefore equals testing for the existence of a long-run relationship. However, the latter concept retains its relevance when some or all of the variables are stationary.

Figure 1 illustrates the concept of cointegration with two simulated nonstationary time series. Despite their stochastic trending behavior, the two processes are bound together long term by a cointegrating relationship. Whenever random shocks drive the processes apart, the y_t process reverts back towards the equilibrium, which is determined by the long-run forcing x_t process. This characterizes an error-correction mechanism. While cointegrated time series are often strongly correlated, such high correlations can also result spuriously by coincidentally similar time trends. In practice, data processes are more intricate than the one depicted in this stylized example. A crucial feature of the ARDL framework therefore is that statistical tests for a systematic long-run relationship can accommodate complex dynamic adjustment processes.[2]

Pesaran and Shin (1998) and Hassler and Wolters (2006) highlight a couple of advantages of the ARDL approach over alternative strategies for cointegration analysis – such as the Engle and Granger (1987) two-step procedure implemented in the community-contributed Stata command `egranger` (Schaffer 2010), or the Phillips and Hansen (1990) fully modified ordinary least squares (FM-OLS) approach implemented in `cointreg` (Wang 2012). First of all, it can accommodate a mixture of stationary and nonstationary variables without the need for pretesting the order of integration. Moreover, the the short-run and long-run coefficients can be consistently estimated in

2. We provide a more detailed explanation of the terminology in Section 2.

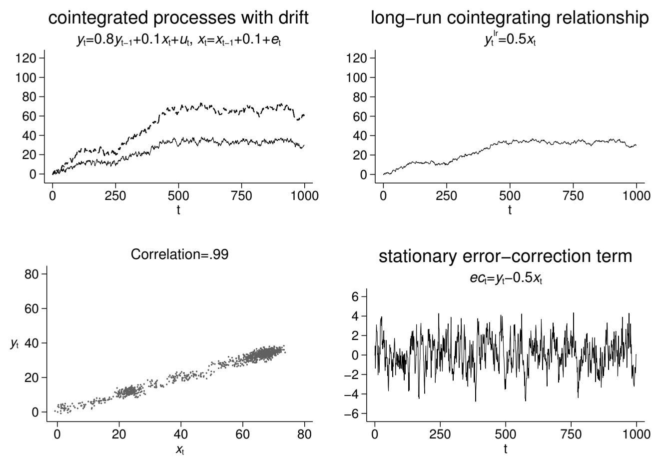

Figure 1: Simulated example of two cointegrated time series

a single step, and the estimator's asymptotic normality eases statistical inference.[3]

Despite these advantages, testing for the existence of a long-run (cointegrating) relationship still requires a bit more effort. The test statistic has a nonstandard distribution which depends on various characteristics of the model and the data, including the integration order of the variables. Pesaran et al. (2001) propose a 'bounds test', which involves comparing the values of conventional F- and t-statistics to pairs of critical values. Outside of these bounds, the test either conclusively rejects or does not reject the null hypothesis. Within the bounds, the test is inconclusive.

In this paper, we present the `ardl` Stata package for the estimation of such single-equation ARDL and EC models. The popular bounds test is implemented as a postestimation feature with recently improved critical value bounds and approximate p-values (Kripfganz and Schneider 2020). Obtained from response-surface regressions using billions of simulated test statistics, these critical values are more precise and exhaustive than earlier critical values tabulated by Pesaran et al. (2001) and Narayan (2005). A key feature of the `ardl` command is the automatic selection of the optimal lag order with the Akaike or Schwarz/Bayesian information criterion. With an increasing number of independent variables, the number of candidate models – which are characterized by all possible combinations of lag orders – quickly is in the tens or even hundreds of thousands. A computationally efficient implementation of this procedure ensures that the optimal model is still found within seconds.

Closely related, Jordan and Philips (2018) recently introduced the `dynardl` Stata

3. Shin et al. (2014) extend the ARDL framework by introducing nonlinearities that allow for asymmetric long-run effects. Such a nonlinear ARDL model can be estimated in Stata using the command `nardl`, implemented by M. Sunder. Here, we restrict our attention to the symmetric case.

command for dynamic simulations of ARDL models. With their `pssbounds` command, they also provide an interface to display the original Pesaran et al. (2001) and Narayan (2005) asymptotic and finite-sample critical values for the bounds test. As we argued above, those critical values are now largely superseded. Moreover, their commands do not perform an automatic lag order selection, which is a key feature of our `ardl` command. Once the optimal model specification is obtained with the `ardl` command, the `dynardl` command can still be a useful complement if a visualization of the dynamic effects is desired.

This article is only concerned with time-series data. For the estimation of ARDL models in a large-T panel-data context, see the community-contributed Stata commands `xtpmg` (Blackburne and Frank 2007), `xtdcce2` (Ditzen 2018, 2021), and `xtivdfreg` (Kripfganz and Sarafidis 2021). The command `xtwest` (Persyn and Westerlund 2008) enables cointegration tests based on panel-data EC models.

In Section 2, we outline the econometric background for the ARDL approach to the analysis of long-run equilibrium relationships; and we provide detailed guidance for the model specification and the bounds test procedure. In Sections 3 and 4, we describe the syntax and options for the `ardl` Stata package. In Section 5, we illustrate the approach by replicating the empirical example of Pesaran et al. (2001). Section 6 concludes.

2 Econometric model and methods

2.1 Spurious regressions and cointegration

Suppose we expect the existence of a relationship between a dependent variable y_t and a set of K explanatory variables $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots, x_{Kt})'$:

$$y_t = b_0 + \mathbf{x}_t'\boldsymbol{\theta} + e_t, \qquad (1)$$

where b_0 is the intercept of the regression line. The data are observed at consecutive time points $t = 1, 2, \dots, T$. We might be tempted to estimate the coefficients $\boldsymbol{\theta}$ by ordinary least squares (OLS). However, when y_t and some or all of the regressors $\mathbf{x}_t$ exhibit a trending behavior over time, it is well known that this can result in spuriously large coefficient estimates, even if there is no underlying relationship among the variables. The coefficient estimates then merely reflect a correlation of the variables that is due to the underlying time trends, instead of a genuine link between changes of $\mathbf{x}_t$ and y_t. A warning sign is often a suspiciously high coefficient of determination (R^2). If the variables are simply fluctuating around a deterministic trend, then including a time trend as an additional regressor – or, equivalently, detrending the data – could provide a remedy:[4]

$$y_t = b_0 + b_1 t + \mathbf{x}_t'\boldsymbol{\theta} + e_t. \qquad (2)$$

However, in contrast to such trend-stationary variables, we often encounter variables which follow a so-called stochastic trend. A simple example is a random-walk process,

4. More generally, a deterministic trend can be any linear or nonlinear function of time t.

$y_t = y_{t-1} + \varepsilon_t$, which has an infinite memory. Any shock to such a process results in a permanent shift in the level of y_t. In other words, the process is not mean reverting. This is a property of the more general class of unit-root processes. To obtain a stationary process that fluctuates around a constant mean, we need to consider the first differences, $\Delta y_t = y_t - y_{t-1}$. Such a unit-root process is said to be integrated of order 1, in short $I(1)$. The integration order indicates the number of times a process needs to be differenced in order to obtain a stationary process. Accordingly, a process that is stationary without further differencing or detrending is labeled $I(0)$. Visualized, a unit-root process typically appears to be locally trending. Therefore, especially when observed for just a relatively short time span, it may be difficult to distinguish it from a trend-stationary process.

Ignoring a deterministic time trend by estimating equation (1) can create an omitted-variable problem. The time trend in the data-generating process (DGP) of y_t becomes part of the error term e_t; and its correlation with the time trend in the DGP of $\mathbf{x}_t$ results in biased estimates. If the trend is stochastic, this will likewise be picked up by the error term; and it could possibly be detected by applying a unit-root test to the regression residuals. The nonstationarity of the error term then implies that the standard asymptotic theory is no longer applicable; and the OLS estimator will have a nondegenerate limiting distribution. In other words, it no longer converges in probability to the true parameter vector $\boldsymbol{\theta}$ but instead to a random variable. Including a deterministic time trend – as in equation (2) – does not help in this case.

Equation 1 can still be a valid regression model if the trending behavior of y_t is actually driven by the trending behavior of $\mathbf{x}_t$. When y_t and some or all of the variables in $\mathbf{x}_t$ are $I(1)$, we say that y_t and the respective regressors are cointegrated if the corresponding coefficients in $\boldsymbol{\theta}$ are nonzero. In that case, the linear combination $y_t - \mathbf{x}_t'\boldsymbol{\theta}$ is stationary, and therefore the same is true for the regression errors. The OLS estimator is super-consistent, as it converges to the true coefficient vector at a faster rate than if the variables were stationary.

Equations (1) and (2) reflect conditional long-run equilibrium relationships – if they exist – to which a process reverts over time. In the short run, the process might divert from this equilibrium, but the above equations are silent about the dynamic evolution of the process when it is off the equilibrium path. Such deviations are transitory, and the elements in the DGP governing them are therefore $I(0)$. These short-run terms are asymptotically negligible relative to the stochastic trend of an $I(1)$ process; their omission does not turn the OLS estimator inconsistent. It is for this reason that an OLS estimation of equation (1) is a valid first step in the Engle and Granger (1987) two-step cointegration test procedure, provided that it has been ascertained with unit-root tests that y_t and $\mathbf{x}_t$ are $I(1)$.[5]

However, the neglected $I(0)$ components in the DGP affect the finite-sample (and possibly the asymptotic) distributions of test statistics and thus invalidate conventional hypothesis tests and regression diagnostics. This is in addition to the spurious-regression

5. For further background reading on the discussed topics in this section, the interested reader is referred to any textbook on time-series econometrics of their choice.

problem when there is no cointegration among the $I(1)$ variables. Therefore, equations (1) and (2) have only limited use for statistical inference on the coefficients $\boldsymbol{\theta}$.

2.2 Autoregressive distributed lag model

To circumvent the problems associated with estimating a static model, we can augment the regression equation with lags of the dependent and independent variables. We can even include another set of L exogenous variables $\mathbf{z}_t$, which may have predictive power to explain the short-term fluctuations of y_t but do not affect its equilibrium path. We assume that all variables in $\mathbf{z}_t$ are stationary. Augmenting the model in this way aims at obtaining a dynamically complete model, in which the regression error term u_t is free of serial correlation:

$$y_t = c_0 + c_1 t + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=0}^{q} \boldsymbol{\beta}_i' \mathbf{x}_{t-i} + \boldsymbol{\gamma}' \mathbf{z}_t + u_t, \tag{3}$$

$t = 1+p^*, \ldots, T$. Leaving aside the variables $\mathbf{z}_t$, this is a general ARDL(p,q,...,q) model with intercept c_0, linear trend $c_1 t$, and lag orders $p \in [1, p^*]$ and $q \in [0, p^*]$.[6] To ensure that there are enough degrees of freedom available to estimate the model's coefficients with sufficient precision, we may need to choose the maximum admissible lag order p^* conservatively. This is especially relevant when the number of observations in the data set (T) is relatively small, and/or the number of variables in $\mathbf{x}_t$ (K) is relatively large.[7]

Given the initial observations $y_1, y_2, \ldots, y_{p^*}$, and the time paths of $\mathbf{x}_t$ and $\mathbf{z}_t$, equation (3) describes the dynamic evolution of y_t over time, irrespective of whether an equilibrium relationship – as postulated in equation (1) or (2) – exists. With such a dynamic model, we no longer need to worry about potentially running a spurious regression. If y_t and $\mathbf{x}_t$ are generated by independent $I(1)$ processes, there now exist combinations of the parameters ϕ_i and $\boldsymbol{\beta}_i$ which still yield stationary regression errors. For example, the model accommodates the case of all variables being independent random walks through the set of coefficients $\phi_1 = 1$, $\phi_i = 0$ for $i > 1$, and $\boldsymbol{\beta}_i = \mathbf{0}$ for all i. The intercept c_0 and the linear time trend $c_1 t$ may or may not be included in the model, depending on the nature of the variables under consideration.[8]

We assume that sufficiently many lags have been included in the ARDL model (3) to purge the error term from any remaining serial correlation and to ensure that the variables $\mathbf{x}_t$ are weakly exogenous/long-run forcing – ruling out any contemporaneous feedback from y_t to $\mathbf{x}_t$. If there exists a stable long-run relationship, conventional

6. Allowing for different lag orders among the components of $\mathbf{x}_t$ is straightforward and can be treated as a special case of the general model by restricting some coefficients to be zero.
7. The importance of the maximum lag order p^* is explained further below. In practice, the data frequency often guides this choice. For instance, it is customary to allow for up to 12 lags with monthly data and up to 4 or 8 lags with quarterly data.
8. In general, other deterministic model components – such as quadratic time trends or impulse dummy variables – can be added. We abstract from them here but note that their inclusion may affect the applicability of the critical values for the bounds test, which is presented further down in this article.

asymptotic theory can be applied for statistical inference on any of the coefficients (Pesaran and Shin 1998). This highlights the importance of testing for the existence of such a long-run relationship, which we consider in Section 2.4.

While the inclusion of further lags improves the regression fit, this comes at the cost of a higher variance of the coefficient estimates. To balance this tradeoff, a data-driven approach to optimal lag selection can be based on the Akaike information criterion (AIC) or the Schwarz/Bayesian information criterion (BIC):

$$\begin{aligned}\text{AIC} &= -2\ln(\mathcal{L}) + 2K^*,\\ \text{BIC} &= -2\ln(\mathcal{L}) + \ln(T^*)K^*,\end{aligned}$$

where $\ln(\mathcal{L})$ is the value of the log-likelihood function from the estimated regression model, $T^* = T - p^*$ is the effective sample size, and $K^* = 2 + p + K(q+1) + L$ is the number of estimated coefficients in model (3). Higher values of $\ln(\mathcal{L})$ indicate a better fit of the model. Thus, we prefer models that deliver a smaller value of the AIC or BIC. However, adding more regressors to the model – in particular, increasing the lag orders p or q – never worsens the fit. To resist the temptation of creating larger and larger models, the model selection criteria contain a penalty term, which is increasing in the number of coefficients (K^*). The BIC has a larger penalty term than the AIC (for $T^* \geq 8$) and therefore tends to select more parsimonious models. The optimal lag orders are then found by estimating model (3) for all possible combinations of p and q, and choosing the model which minimizes the AIC or BIC.

For the comparability of the model selection criteria, it is imperative that we base all regressions on the same estimation sample. This is the reason for initially choosing a fixed maximum lag order p^*. When both p and q are smaller than p^*, the estimation of model (3) does not use all of the available observations. This is the price we need to pay for consulting the model selection criteria. Once the optimal lag orders p and q have been found, we can subsequently re-estimate the model, utilizing all available observations by setting $p^* = \max(p, q)$.

2.3 Error-correction representation

The coefficients in the ARDL model (3) have a less straightforward interpretation than those in the static model. To regain a better interpretability, we can reformulate the ARDL model in EC representation (Hassler and Wolters 2006):[9]

$$\Delta y_t = c_0 + c_1 t - \alpha(y_{t-1} - \boldsymbol{\theta}\mathbf{x}_{t-1}) + \sum_{i=1}^{p-1}\psi_{yi}\Delta y_{t-i} + \boldsymbol{\omega}'\Delta\mathbf{x}_t + \sum_{i=1}^{q-1}\boldsymbol{\psi}'_{xi}\Delta\mathbf{x}_{t-i} + \boldsymbol{\gamma}'\mathbf{z}_t + u_t. \quad (4)$$

9. By convention, the summations evaluate to zero if the upper limit is smaller than the lower limit.

The coefficients in equation (4) can be mapped in a straightforward algebraic way to the coefficients in equation (3):

$$\alpha = 1 - \sum_{i=1}^{p} \phi_i, \qquad \boldsymbol{\theta} = \frac{\sum_{j=0}^{q} \boldsymbol{\beta}_j}{\alpha},$$

$$\psi_{yi} = - \sum_{j=i+1}^{p} \phi_j, \qquad \boldsymbol{\omega} = \boldsymbol{\beta}_0, \qquad \boldsymbol{\psi}_{xi} = - \sum_{j=i+1}^{q} \boldsymbol{\beta}_j.$$

Now recall the hypothesized long-run equilibrium relationship between y_t and $\mathbf{x}_t$ in equation (1) or (2). For the moment ignoring the intercept and linear time trend, the deviations from this equilibrium, $e_{t-1} = y_{t-1} - \boldsymbol{\theta}\mathbf{x}_{t-1}$, can be found again in the EC model (4). Due to the nonlinear interaction between the coefficients α and $\boldsymbol{\theta}$, we cannot directly estimate equation (4) by OLS. However, given the mapping above, we can recover consistent estimates of all coefficients from estimating the ARDL model (3). Yet, a computationally more convenient approach is to instead estimate the following model:[10]

$$\Delta y_t = c_0 + c_1 t + \pi_y y_{t-1} + \boldsymbol{\pi}_x \mathbf{x}_{t-1} + \sum_{i=1}^{p-1} \psi_{yi} \Delta y_{t-i} + \boldsymbol{\omega}' \Delta \mathbf{x}_t + \sum_{i=1}^{q-1} \boldsymbol{\psi}'_{xi} \Delta \mathbf{x}_{t-i} + \boldsymbol{\gamma}' \mathbf{z}_t + u_t, \quad (5)$$

from which we can easily recover the so-called speed-of-adjustment coefficient $\alpha = -\pi_y$ and the long-run coefficients $\boldsymbol{\theta} = \boldsymbol{\pi}_x / \alpha$. The corresponding standard errors can be computed with the delta method (Pesaran and Shin 1998). Notice that equation (5) collapses to the well-known augmented Dickey and Fuller (1979) regression for unit-root testing when no explanatory variables $\mathbf{x}_t$ and $\mathbf{z}_t$ are present ($K = L = 0$).

An important role is played by the speed-of-adjustment coefficient α, which is the coefficient (with opposite sign) of the EC term e_{t-1}. It tells us how fast the process for y_t reverts back to its long-run relationship when this equilibrium is distorted. $\alpha = 1$ would imply that – in the absence of any other short-run fluctuations – any deviation from the equilibrium is fully corrected immediately in the period after the distortion occurs. In contrast, $\alpha = 0$ would imply that the process never returns back to its equilibrium path. Values of α between these two boundaries reflect a partial-adjustment process, where the gap to the equilibrium is gradually closed over time.[11]

Clearly, $\boldsymbol{\theta} \neq \mathbf{0}$ is not a sufficient condition for the existence of a conditional long-run relationship between the levels of y_t and $\mathbf{x}_t$. When $\alpha = 0$, then y_t is $I(1)$ and no such relationship exists. In the opposite scenario, when $\boldsymbol{\theta} = \mathbf{0}$ and $\alpha \in (0, 2)$, then y_t is (trend) stationary, irrespective of the integration order of the components in $\mathbf{x}_t$. For a long-run level relationship to exist, we need both $\boldsymbol{\theta} \neq \mathbf{0}$ and $\alpha \in (0, 2)$. In

10. When called with the option `ec1`, the `ardl` command estimates equation (5) but reports the coefficients for equation (4).
11. While in the following we allow α to fall into the interval $[0, 2)$, we do not pay particular attention to the oscillating/overshooting case $\alpha > 1$ in this paper. We also rule out explosive processes, which result under $\alpha < 0$. An estimate of α outside of the reasonable region $[0, 1]$ should be seen as a warning signal for potential model misspecification.

this case – as long as the elements of $\mathbf{x}_t$ are not cointegrated among themselves – the integration properties of $\mathbf{x}_t$ determine the integration order of y_t. If the variables in $\mathbf{x}_t$ with nonzero long-run coefficient are $I(1)$, then y_t is $I(1)$ as well, and the conditional long-run relationship corresponds to a cointegrating relationship.

In this context, notice that the assumption of $\mathbf{x}_t$ being long-run forcing for y_t implies that there can exist at most one cointegrating relationship which involves y_t. This does not rule out further cointegrating relationships among the elements of $\mathbf{x}_t$. Thus, without further inspection, a cointegration rank larger than 1 for the entire system $(y_t, \mathbf{x}_t')'$ does not necessarily imply a violation of this assumption. However, if there is reason to suspect multiple cointegrating relationships involving y_t, then a single-equation ARDL/EC model is inappropriate.[12] Instead, this would call for the estimation of a VAR model or – analogously to the EC representation of the ARDL model – a vector error-correction (VEC) model.[13]

The remaining coefficients ψ_{yi}, $\boldsymbol{\omega}$, $\boldsymbol{\psi}_{xi}$, and $\boldsymbol{\gamma}$ in equation (4) capture the short-run dynamics that are not prescribed by the equilibrium-reverting forces.[14] They are not only relevant for making dynamic forecasts, but also play a role for choosing appropriate critical values when testing for the existence of a long-run relationship, which we explore in Section 2.4.

A complication arises if $q = 0$ for some or all of the long-run forcing variables. In that situation, $\boldsymbol{\pi}_x = \boldsymbol{\omega}$, which implies that the corresponding variance-covariance matrix of the coefficient estimates in equation (5) is rank deficient. To avoid this complication, the EC representation can be equivalently formulated with the levels of the long-run forcing variables expressed in period t instead of $t-1$:

$$\Delta y_t = c_0 + c_1 t - \alpha(y_{t-1} - \boldsymbol{\theta}\mathbf{x}_t) + \sum_{i=1}^{p-1} \psi_{yi} \Delta y_{t-i} + \sum_{i=0}^{q-1} \boldsymbol{\psi}_{xi}' \Delta \mathbf{x}_{t-i} + \boldsymbol{\gamma}'\mathbf{z}_t + u_t, \tag{6}$$

with the same parameter restrictions as defined above. Notice that $\boldsymbol{\omega}'\Delta\mathbf{x}_t$ is replaced by $\boldsymbol{\psi}_{x0}'\Delta\mathbf{x}_t$. The interpretation of the long-run coefficients $\boldsymbol{\theta}$ does not change because the time subscript does not matter when the process is in equilibrium. The equation to be estimated in this case becomes

$$\Delta y_t = c_0 + c_1 t + \pi_y y_{t-1} + \boldsymbol{\pi}_x \mathbf{x}_t + \sum_{i=1}^{p-1} \psi_{yi} \Delta y_{t-i} + \sum_{i=0}^{q-1} \boldsymbol{\psi}_{xi}' \Delta \mathbf{x}_{t-i} + \boldsymbol{\gamma}'\mathbf{z}_t + u_t, \tag{7}$$

where the coefficients $\boldsymbol{\pi}_x$ are identical to the corresponding coefficients in equation (5), despite the change in the time subscript.[15]

12. As a consequence, it is not permissible to run several ARDL regressions involving the variables $(y_t, \mathbf{x}_t')'$, in which the dependent variable of one regression becomes an independent variable in other regressions.
13. See `var`, `vec`, and related Stata commands.
14. Strictly speaking, the error correction governed by the coefficient α is a short-run adjustment as well.
15. When called with the option `ec`, the `ardl` command estimates equation (7) but reports the coeffi-

2.4 Bounds test

Even though we can consistently estimate all coefficients in the ARDL regression model (3) – or its EC representations (4) and (6) – testing for the existence of a long-run relationship involves a bit more effort. This is because the process for y_t contains a unit root under the null hypothesis of no long-run relationship, and therefore the test statistics have nonstandard distributions. Moreover, we need to combine the evidence from multiple tests because each of them only provides a partial picture. The testing procedure has up to three steps:

1. First, we test the joint null hypothesis

   $$H_0 : (\pi_y = 0) \cap (\boldsymbol{\pi}_x = \mathbf{0})$$

   versus the alternative hypothesis

   $$H_1 : (\pi_y \neq 0) \cup (\boldsymbol{\pi}_x \neq \mathbf{0}).$$

   The hypotheses are not directly formulated in terms of the long-run coefficients $\boldsymbol{\theta}$, because they are not well defined when $\pi_y = 0$. Instead, the test is formulated as a test for valid exclusion of the level terms y_{t-1} and $\mathbf{x}_{t-1}$ (or $\mathbf{x}_t$) in equation (5) or (7). The test statistic is a conventional F-statistic for joint validity of the $K + 1$ restrictions imposed under the null hypothesis. However, the nonstandard distribution requires the uses of different critical values, which we discuss further below. If the null hypothesis is not rejected, we conclude that there is no statistical evidence in favor of a long-run level relationship between y_t and $\mathbf{x}_t$. Otherwise, we should proceed with the following steps due to the possibility of degenerate cases, which are not ruled out by the alternative hypothesis of this first step.

2. If the null hypothesis from step 1 is rejected, we need to rule out the special case that y_t is $I(1)$ but not cointegrated with any variable in $\mathbf{x}_t$. This is done by testing

   $$H_0 : \pi_y = 0 \quad \text{versus} \quad H_1 : \pi_y < 0.$$

   The test statistic is a conventional t-statistic for the statistical insignificance of the negative speed-of-adjustment estimate with a one-sided rejection region. As in step 1, the distribution of the test statistic is nonstandard and the usual critical values do not apply. If the null hypothesis is not rejected, we conclude again that there is no statistical evidence of a long-run level relationship. Otherwise, we proceed with step 3.

3. If the null hypotheses in steps 1 and 2 are both rejected, we eventually consider the degenerate case that y_t is (trend) stationary, but not part of a long-run relationship

cients for equation (6). As an aside, when option `ec1` is specified – which normally commands time subscripts $t - 1$ for the long-run forcing variables – a subscript t is used in the estimation equation for those variables whose lag order is$q = 0$. The reported results are still reparameterized as in equation (4), however, incorporating the constraint on $\boldsymbol{\omega}$.

with $\mathbf{x}_t$. For this purpose, we can use conventional Wald tests for the joint (or individual) statistical insignificance of the long-run coefficients:

$$H_0 : \boldsymbol{\theta} = \mathbf{0} \quad \text{versus} \quad H_1 : \boldsymbol{\theta} \neq \mathbf{0}.$$

We base this test on the long-run coefficients $\boldsymbol{\theta}$ rather than $\boldsymbol{\pi}_x$ because the OLS estimator of $\boldsymbol{\theta}$ has an asymptotic normal distribution (Pesaran and Shin 1998), irrespective of the integration orders of the variables in $\mathbf{x}_t$, assuming that $\alpha => 0$ as indicated by the test result from step 2. Thus, the conventional critical values can be used.

The rejection of the null hypotheses from all three steps is necessary to conclude that there is statistical evidence in favor of a long-run relationship; that is, $(\alpha > 0) \cap (\boldsymbol{\theta} \neq \mathbf{0})$. It is clear that the alternative hypothesis in step 1 does not rule out the two degenerate cases, which are the subject of steps 2 and 3. Yet, we should still start with step 1 because it is carried out under less restrictive assumptions on the DGP than step 2.[16]

For the test statistics in steps 1 and 2, Pesaran et al. (2001) derive the asymptotic distributions under two scenarios. In the first scenario, all long-run forcing variables $\mathbf{x}_t$ are individually $I(0)$. In the second scenario, all of them are $I(1)$ and not mutually cointegrated. When the (co-)integration properties of $\mathbf{x}_t$ are unknown, the corresponding critical values form lower and upper bounds, respectively. Conclusive evidence is possible when the value of the test statistic falls outside of these bounds. The region for not rejecting the null hypothesis is below the lower bound (closer to zero), and the rejection region is above the upper bound. The test is inconclusive if the test statistic falls between the two bounds. Because the distributions have a nonstandard form, critical values have to be obtained by simulations. This is complicated by the fact that the distributions depend on the number of variables in $\mathbf{x}_t$. For $K \leq 10$, Pesaran et al. (2001) tabulated near-asymptotic critical values for the F-statistic in step 1 and the t-statistic in step 2. However, the asymptotic distributions might be poor approximations when the sample size is relatively small.

It is important to keep in mind that the distributions and critical values are obtained under the assumption of independent and identically normally distributed errors u_t. As mentioned earlier, a standard procedure for dealing with suspected serial correlation is to increase the lag orders p and/or q in the ARDL model. While the $p + Kq$ short-run terms in the EC representation do not affect the asymptotic distributions of the test statistics, they are relevant for the finite-sample distributions. Consequently, different critical values are needed for each combination of T^*, K, and $p + Kq$, separately for the lower and upper bound. Instead of tabulating vast amounts of critical values, Kripfganz and Schneider (2020) estimated response-surface regressions, which can predict critical values for any desired sample size, number of long-run forcing variables, and lag order. This includes asymptotic critical values. Another important advantage of this approach is the ability to compute approximate p-values, which facilitate statistical inference.

As an additional dimension, the distributions of the test statistics – and consequently

16. For technical details and a full set of assumptions, see Pesaran et al. (2001).

the critical values – also depend on the choice of deterministic model components. In the ARDL model (3) – and its EC representations (4) and (6) – we have allowed for an intercept c_0 and a linear time trend $c_1 t$. We can distinguish the following five cases:

1. No deterministic model components are included, $c_0 = c_1 = 0$.
2. A restricted intercept is included, $c_0 = \alpha b_0$, but no time trend, $c_1 = 0$.
3. An unrestricted intercept is included, $c_0 \neq 0$, but no time trend, $c_1 = 0$.
4. An unrestricted intercept is included, $c_0 \neq 0$, and a restricted time trend, $c_1 = \alpha b_1$.
5. Both deterministic model components are unrestricted, $c_0 \neq 0$ and $c_1 \neq 0$.

A decision about the relevant case can often be guided by a visual inspection of the time series. Cases 1 and 2 are in line with a process y_t which could reasonably be an $I(1)$ process without drift under the null hypothesis of no long-run level relationship. Under the alternative hypothesis, y_t would either be $I(0)$ or cointegrated with $\mathbf{x}_t$. Case 1 is most appropriate if y_t and $\mathbf{x}_t$ fluctuate around a zero mean, or if any nonzero means cancel out in the long-run level relationship; that is, $b_0 = 0$ in equation (1). The latter condition is hard to verify ex ante, such that case 2 is often the safer option whenever some variables have a nonzero mean.

If y_t appears to be trending, it could be an $I(1)$ process with drift under the null hypothesis. This calls for case 3 or 4. Under the alternative hypothesis, y_t would either be trend stationary or cointegrated with $\mathbf{x}_t$. Case 3 is most appropriate if the trend in y_t is entirely attributable to a trend in $\mathbf{x}_t$; that is, $b_1 = 0$ in equation (2). Again, this may be difficult to justify ex ante. Despite the fact that case 3 is most commonly applied in the empirical practice, case 4 is generally the safer option when there is insufficient knowledge about the source of the observed time trend.[17]

Especially when the sample size is relatively small, it might be difficult to distinguish visually between a mildly drifting unit-root process under the null hypothesis and a stationary process which is fluctuating around a constant mean under the alternative hypothesis. This can be another relevant situation for case 3. Similarly, case 5 could be used to statistically discriminate between a unit-root process with faster – although hardly noticeable – than linear growth (or decline) and a trend-stationary process. For most practical applications, this might be a rather irrelevant scenario.

Notice that the restrictions on the intercept or linear trend under cases 2 and 4 do not affect the estimation of the ARDL model because it is irrelevant whether we treat c_0 (c_1) or b_0 (b_1) as a free parameter to be estimated. Under case 1, equation (3) is estimated without intercept and trend. Under cases 2 and 3, an intercept is included in the regression. Under cases 4 and 5, an intercept and linear time trend are included. However, the restrictions are incorporated into step 1 of the bounds test.

17. Recall the discussion in Section 2.1 about including a linear time trend in the equilibrium relationship.

The null hypothesis implies $c_0 = 0$ under case 2 and $c_1 = 0$ under case 4, which adds an additional linear restriction to the test. Consequently, different critical values apply. Because these additional restrictions do not alter the underlying DGP, the critical values for the single-hypothesis test in step 2 are the same for cases 2 and 3, and similarly for cases 4 and 5.[18]

The following stages characterize a stylized ARDL approach to testing for the existence of a conditional long-run level relationship:

1. Decide about the candidate variables $\mathbf{x}_t$ that are assumed to be long-run forcing for y_t. These variables can be either $I(0)$ or $I(1)$. No pretesting is necessary, unless we suspect that a variable might be $I(2)$. Stationary variables $\mathbf{z}_t$ that are suspected to affect the short-run dynamics – but not the long-run equilibrium – can be added to the ARDL model as well. If there is doubt about the stationarity of $\mathbf{z}_t$, unit-root tests can be carried out.

2. Decide about the deterministic model components to be included in the model, and whether the constant or linear trend coefficient should be restricted; that is, choose one of the 5 cases above.

3. Choose a maximum lag order p^*, ensuring that sufficiently many degrees of freedom are available for the estimation of the model parameters.[19] Keeping the estimation sample fixed, use the AIC or BIC to obtain the optimal lag orders p and q. To assert that the model is dynamically complete, a serial-correlation test could be of assistance. If there is concern about remaining serial correlation, the AIC might be preferred over the BIC, because it tends to select less parsimonious models. Additional specification tests – for example, tests for heteroskedasticity and normality of the errors – could be used to check whether the assumptions underlying the bounds test are met.

4. If $\max(p, q) < p^*$, re-estimate the ARDL($p, q, \dots, q$) model, now using all available observations. Use the EC representation to check the plausibility of the coefficient estimates. For example, an implausible estimate of α, which is clearly outside of the interval $[0, 2)$, might give rise to concern about the correct model specification or a potential overparameterization of the model.

5. Follow the three steps of the bound test procedure. For steps 1 and 2, do not reject the null hypothesis if the value of the test statistic is below – that is, closer to zero – the lower bound of the Kripfganz and Schneider (2020) critical values. Reject the null hypothesis (and proceed with the next testing step) if the test statistic exceeds the upper-bound critical value.

6. If there is conclusive statistical evidence in favor of a long-run relationship, consider re-estimating a more parsimonious model with lag orders selected by the BIC.

18. See again Pesaran et al. (2001) for further discussion.
19. The Kripfganz and Schneider (2020) critical values are only available if there are at least twice as many observations T^* than coefficients K^*. For reliable inference, a much higher ratio is usually recommended.

If there is evidence against a long-run level relationship, consider re-estimating an ARDL model in first differences to obtain more efficient estimates:

$$\Delta y_t = c_0 + c_1 t + \sum_{i=1}^{p-1} \psi_{yi} \Delta y_{t-i} + \sum_{i=0}^{q-1} \boldsymbol{\psi}'_{xi} \Delta \mathbf{x}_{t-i} + \boldsymbol{\gamma}' \mathbf{z}_t + u_t, \tag{8}$$

which is a restricted version of equation (7) with $\pi_y = 0$ and $\boldsymbol{\pi}_x = \mathbf{0}$. In both cases, it might be worth removing variables which do not help to improve the model fit. This re-estimation stage can be skipped if there is no interest in further statistical analysis – for example, forecasting – beyond the exploration of a level relationship.

In order to avoid pretesting problems, model simplifications – like those at stage 6 – should be kept to a minimum before the bounds test is performed. Also note that there is no need to separately estimate a static model in levels if the bounds test provides evidence in favor of a long-run relationship. As discussed earlier, the respective long-run coefficients can be inferred directly from the EC representation (4) or (6).

3 The `ardl` command

3.1 Syntax

`ardl` *depvar* [*indepvars*] [*if*] [*in*] [, `lags(`*numlist*`)` `exog(`*varlist*`)` `ec` `ec1` `noconstant` `trendvar`[`(`*varname*`)`] `restricted` `regstore(`*name*`)` `perfect` `maxlags(`*numlist*`)` `aic` `bic` `maxcombs(`*#*`)` `matcrit(`*name*`)` `nofast` `dots` *display_options*]

3.2 Options

`lags(`*numlist*`)` specifies the number of lags for some or all regressors. The first number specifies the lag length p for *depvar* (y_t), which has to be larger than zero. The following numbers specify the lag lengths q for the independent variables in the order they appear in *indepvars* ($\mathbf{x}_t$), which can be zero or higher. Missing values are allowed; they indicate that the respective lag order is not prespecified but instead determined with information criteria. If *numlist* contains only one element, the same lag order is applied to all variables. Otherwise, the number of elements in *numlist* must equal the number of variables in *depvar* and *indepvars*.

`exog(`*varlist*`)` specifies additional variables ($\mathbf{z}_t$) to be added to regression. An automatic lag order selection is not performed for these variables.

`ec` requests to display the results in error-correction form. *indepvars* enter the long-run relationship with time subscript t, as in equation (6).

`ec1` requests to display the results in error-correction form. *indepvars* enter the long-run relationship with time subscript $t-1$, as in equation (4).

`noconstant` supresses the constant term. Specifying this option implies that the bounds test uses critical values for case 1.

`trendvar`[`(`*varname*`)`] specifies a linear time trend to be added to the regression. *varname* must be a variable that is collinear with *timevar*, the variable which is used with `tsset` to declare the data to be time-series data. Specifying `trendvar` is equivalent to `trendvar(`*timevar*`)`. Specifying this option implies that the bounds test uses critical values for case 4 or 5.

`restricted` specifies that the constant term or the time trend, if specified, will be restricted for the purpose of the bounds test. The restricted deterministic component will be displayed in the long-run section of the error-correction output. Specifying this option implies that the bounds test uses critical values for case 2 or 4.

`regstore(`*name*`)` requests to store the estimation results from the underlying `regress` command. These are the OLS estimates of equations (5) or (7) when option `ec1` or `ec0` is specified, respectively, and equation (3) otherwise.

`perfect` omits the collinearity check among the regressors.

`maxlags(`*numlist*`)` specifies the maximum lag order p^* for the optimal lag order selection. The first number specifies the maximum lag length for *depvar* (y_t), which has to be larger than zero. The following numbers specify the maximum lag lengths for the independent variables in the order they appear in *indepvars* ($\mathbf{x}_t$), which can be zero or higher. Missing values are allowed; they indicate that the default maximum lag order 4 is to be used. If *numlist* contains only one element, the same maximum lag order is applied to all variables. Otherwise, the number of elements in *numlist* must equal the number of variables in *depvar* and *indepvars*.

`aic` requests the optimal lag lengths to be determined with the Akaike information criterion.

`bic`, the default, requests the optimal lag lengths to be determined with the Bayesian information criterion.

`maxcombs(#)` restricts the maximum number of lag permutations for the automatic lag selection. The default is 100,000, or 500 if option `nofast` is specified. Higher values are possible.[20]

`matcrit(`*name*`)` requests to save the lag permutations and the respective information criterion in a matrix named *name*.

`nofast` uses the `regress` command instead of dedicated Mata code to run the auxiliary regressions for the optimal lag order selection. This is much slower but might be numerically more robust in rare cases.

---

20. The purpose of this option is to prevent the optimal lag order selection from taking a lot of time without explicit user consent.

`dots` displays a progress bar for the optimal lag order selection. This is useful when there are many permutations, due to a large number of variables and high maximum lag orders. Each dot represents a 1% progress in the evaluation of candidate models.

display_options: `noctable`, `noheader`, `noomitted`, `vsquish`, `nolstretch`, `cformat(`*%fmt*`)`, `pformat(`*%fmt*`)`, `sformat(`*%fmt*`)`.

3.3 Saved results

`ardl` stores the following results in `e()`:[21]

Scalars			
`e(N)`	number of observations	`e(rank)`	rank of `e(V)`
`e(df_m)`	model degrees of freedom	`e(F)`	F-statistic *
`e(df_r)`	residual degrees of freedom	`e(case)`	case number for deterministic components *
`e(mss)`	model sum of squares		
`e(rss)`	residual sum of squares	`e(F_pss)`	bounds test F-statistic *
`e(rmse)`	root mean square error	`e(t_pss)`	bounds test t-statistic *
`e(r2)`	R-squared	`e(numcombs)`	number of lag combinations *
`e(r2_a)`	adjusted R-squared	`e(N_gaps)`	number of gaps in sample
`e(ll)`	log likelihood under assumption of i.i.d. normal errors	`e(tmin)`	first time period in sample
		`e(tmax)`	last time period in sample
`e(ll_0)`	log-likelihood, constant only		
Macros			
`e(cmd)`	`ardl`	`e(tmaxs)`	formatted maximum time
`e(cmdline)`	command as typed	`e(regressors)`	full set of regressors
`e(cmdversion)`	version of the `ardl` command	`e(det)`	deterministic components *
`e(model)`	`level` or `ec`	`e(exogvars)`	exogenous variables *
`e(title)`	title in estimation output	`e(srvars)`	short-run regressors *
`e(estat_cmd)`	`ardl_estat`	`e(lrdet)`	long-run deterministic component *
`e(predict)`	`ardl_p`		
`e(tsfmt)`	format for the time variable	`e(lrxvars)`	long-run regressors *
`e(tvar)`	time variable	`e(properties)`	`b V`
`e(tmins)`	formatted minimum time	`e(depvar)`	name of dependent variable
Matrices			
`e(b)`	coefficient vector	`e(maxlags)`	maximum lag lengths
`e(V)`	variance-covariance matrix	`e(lags)`	lag lengths in ARDL model
Functions			
`e(sample)`	marks estimation sample		

4 Postestimation commands

A large number of standard postestimation commands for the `regress` command can be used after the `ardl` command. Importantly, the results obtained with some of them can differ depending on whether the model is specified in the ARDL level form (3) or one of the EC forms (4) or (6). For example, the `estat ovtest` includes higher-order powers of the dependent variable – which is either y_t or Δy_t – as regressors in an auxiliary regression. This complication does not apply to postestimation commands based on residuals – such as `estat bgodfrey` and `estat imtest` – because the error term u_t is

21. Starred results (*) are not always saved.

unaffected by the model's reparameterization.

The Pesaran et al. (2001) bounds test for the existence of a long-run level relationship with Kripfganz and Schneider (2020) critical values and approximate p-values – as discussed in Section 2.4 – is implemented in the postestimation command `estat ectest`. It requires the option `ec` or `ec1` to be specified with the `ardl` command.

4.1 Syntax

`estat ectest` [, `siglevels(`*numlist*`)` `asymptotic` `nocritval` `norule` `nodecision`]

4.2 Options

`siglevels(`*numlist*`)` requests to show critical values for all significance levels in *numlist*. The default is `siglevels(10 5 1)`. The admissible significance levels are any subset of the Stata number list `0.01 0.02 0.05 0.10(0.10)0.90 1.00(0.50)98.50 99.00(0.10)99.90 99.95 99.98 99.99`.

`asymptotic` requests that the actual sample size be ignored and asymptotic critical values be shown instead.

`nocritval` suppresses display of the critical values table.

`norule` suppresses display of the decision rule.

`nodecision` suppresses display of the decision table.

5 Example

To illustrate the `ardl` command, we replicate the empirical analysis of Pesaran et al. (2001). They estimate an earnings equation with macroeconomic data for the United Kingdom. The raw data are not in Stata format but can be imported as follows:[22]

```
. infile double (ERPR UDEN GDPMS FCAS CGGS PBRENT YOG RXD GDPA CGG TYEM EENIC WFP
> YPROM YMF EMF ENMF EMPNIC NIS OCR ILOU WFEMP RPIX PYNONG TE TD) using earn2.dat
(116 observations read)
. replace ILOU = . if ILOU > 99999999
(4 real changes made, 4 to missing)
. range qtr tq(1969q1) tq(1997q4)
. format qtr %tq
. quietly compress
```

22. The data set and a brief description of the variables can be downloaded from the Journal of Applied Econometrics data archive: http://qed.econ.queensu.ca/jae/2001-v16.3/pesaran-shin-smith/. There are two data files in the archive. Because the variable for the real wage in the file `earn1.dat` – which contains the regression variables – does not match its definition, we reconstruct all variables from the file `earn2.dat` – which contains the raw data.

```
. tsset qtr
Time variable: qtr, 1969q1 to 1997q4
Delta: 1 quarter
```

The dependent variable y_t is the real wage (`w`). The presumed long-run forcing variables $\mathbf{x}_t$ are the labor productivity (`Prod`), the unemployment rate (`UR`), a wedge between the real wage from the firm's perspective and the real wage from the union's perspective (`Wedge`), and a measure of union power (`Union`). We construct these main variables from the raw data:

```
. generate w = ln(ERPR/PYNONG)
. generate Prod = ln((YPROM+278.29*YMF)/(EMF+ENMF))
. generate UR = ln(100*ILOU/(ILOU+WFEMP))
(4 missing values generated)
. generate Wedge = ln(1+TE) + ln(1-TD) - ln(RPIX/PYNONG)
. generate Union = ln(UDEN)
```

Because the data for `UR` are not available before 1970, the first 4 quarters in the data set are not relevant for the subsequent analysis. The subsequent 8 quarters (1970–1971) are initially set aside for the optimal lag selection. Similar to Pesaran et al. (2001, Figures 1–3), we graph the levels of these variables in Figure 2 for a visual inspection:

```
. generate byte smpl = inrange(qtr, tq(1972q1), tq(1997q4))
. set scheme sj
. tsline w Prod UR if smpl, ylabel(0(0.5)4, format(%2.1f) angle(horizontal))
> tlabel(1972q1(8)1997q4, format(%tqCCYY)) xtitle("") legend(rows(1)) name(ts1)
. tsline Wedge Union if smpl, ylabel(-0.8(0.1)0, format(%2.1f) angle(horizontal))
> tlabel(1972q1(8)1997q4, format(%tqCCYY)) xtitle("") name(ts2)
. graph combine ts1 ts2, ysize(2) xsize(5)
```

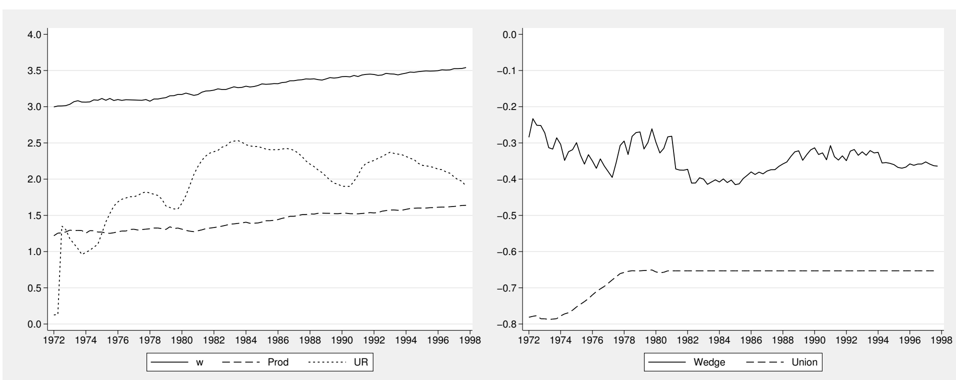

Figure 2: Time-series graphs for the main regression variables

At first glance, the `Wedge` appears to be the only variable which could possibly be stationary over the whole sample period. For `w`, `Prod`, and `UR`, it is not immediately obvious whether they are generated by unit-root or trend-stationary processes. Because the ARDL approach can deal with both $I(0)$ and $I(1)$ variables, we refrain from unit-root pretesting. Even if some of the variables are driven by a linear time trend, it is not a priori clear whether a deterministic trend needs to be included in the regression model. It could well be that the time trend underlying labor productivity fully explains the time trend of real wages.

`Union` could be a stationary variable once it is adjusted for the level shift in the 1970s. The observed change in the union power is likely the result of specific income policies during the periods 1974–1975 and 1975–1979. Pesaran et al. (2001) account for them with two dummy variables:

```
. generate byte D7475 = inrange(qtr, tq(1974q1), tq(1975q4))
. generate byte D7579 = inrange(qtr, tq(1975q1), tq(1979q4))
```

For ease of exposition, we treat these dummy variables as additional short-run terms $\mathbf{z}_t$. However, strictly speaking, such 'one-off' dummy variables are not $I(0)$ variables. They are deterministic model components. As pointed out by Pesaran et al. (2001), the asymptotic theory for the bounds test is unaffected by the inclusion of such variables as long as the fraction of time periods covered by them tends to zero as $T \to \infty$. Finite-sample critical values covering this particular type of deterministic components are not readily available. The subsequent results are therefore to be taken with a grain of salt, although the sample size is reasonably large to feel comfortable about ignoring this complication.

To choose an optimal lag order for the purpose of the bounds test, Pesaran et al. (2001) initially fit models with a constant lag order $p = q \in [1, 7]$ for all variables besides `Prod`, for which they set $q = 1$ after observing that its lagged differences do not have a statistically significant contribution. They fit each of these models separately with and without a deterministic trend. These ARDL(p, 1, p, p, p) models can be easily estimated with the `ardl` command by specifying the lag orders with the `lags()` option. A linear time trend is included with the `trendvar` option. The AIC and BIC can be displayed with the `estat ic` postestimation command. Pesaran et al. (2001) also inspect the LM test for no serial correlation of the residuals against orders 1 and 4, which is a critical assumption for the bounds test. It comes in handy that most standard postestimation commands for the `regress` command are also available after `ardl`.

```
. forvalues p = 1 / 7 {
  2.         quietly ardl w Prod UR Wedge Union if smpl, exog(D7475 D7579)
> lags(`p´ 1 `p´ `p´ `p´) trendvar
  3.         estat ic
  4.         estat bgodfrey, lags(1 4)
  5.         quietly ardl w Prod UR Wedge Union if smpl, exog(D7475 D7579)
> lags(`p´ 1 `p´ `p´ `p´)
  6.         estat ic
  7.         estat bgodfrey, lags(1 4)
  8. }
  (output omitted)
```

The results replicate those in Pesaran et al. (2001, Table I),[23] with the exception of the very last LM test statistic against the alternative of a fourth-order serial correlation for the model with $p = 7$ and no time trend. Pesaran et al. (2001) erroneously report the p-value of 0.64 instead of the value of the test statistic, which is 2.51. Among these models, the smallest value of the AIC is attained with $p = 6$, irrespective of the trend specification. With the BIC, the preferred model has $p = 1$ when a trend is included, and $p = 4$ otherwise. The serial correlation tests do not provide comfort for lag orders $p \leq 4$. For the bounds test, it would thus be advisable to base the lag order decision on the AIC.

Under the null hypothesis of no long-run relationship, the level of `w` is not determined by the levels of the independent variables – especially `Prod`, which shows a similar trending behavior. Thus, the observed time trend of `w` can only be a result of a unit-root process with positive drift (when $\alpha = 0$) or a trend-stationary process (when $\alpha > 0$ but $\boldsymbol{\theta} = \mathbf{0}$). Recalling our discussion in Section 2.4, this calls for either a model without trend but unrestricted intercept (case 3), or a model with restricted time trend (case 4). Case 3 can be justified on the grounds that the trend of `w` is likely explained by the trend of `Prod` when a long-run relationship among them exists. This choice is further supported by the observation that the time trend is statistically insignificant in the models estimated above, with the exception of $p = 1$. However, the AIC still slightly prefers the model with trend under the optimal lag order $p = 6$.

Avoiding a final verdict about the time trend, let us conduct the bounds test both under case 3 and 4. We do this by re-estimating the chosen model ($p = 6$) in EC representation with the `ec` (or `ec1`) option. For case 3, we do not specify a time trend. For case 4, we specify the option `restricted` in addition to the time trend. Subsequently, we obtain the bounds test results with the postestimation command `estat ectest`:

```
. quietly ardl w Prod UR Wedge Union if smpl, exog(D7475 D7579) lags(6 1 6 6 6) ec
. estat ectest
Pesaran, Shin, and Smith (2001) bounds test
H0: no level relationship                                     F =     5.421
Case 3                                                        t =    -3.475
Finite sample (4 variables, 104 observations, 26 short-run coefficients)
Kripfganz and Schneider (2020) critical values and approximate p-values
   | 10%              | 5%               | 1%               | p-value
   |    I(0)     I(1) |    I(0)     I(1) |    I(0)     I(1) |    I(0)     I(1)
---+------------------+------------------+------------------+-----------------
 F |   2.362    3.646 |   2.806    4.226 |   3.800    5.502 |   0.001    0.011
 t |  -2.447   -3.499 |  -2.777   -3.873 |  -3.421   -4.589 |   0.009    0.104
do not reject H0 if
    either F or t are closer to zero than critical values for I(0) variables
      (if either p-value  > desired level for I(0) variables)
reject H0 if
    both F and t are more extreme than critical values for I(1) variables
      (if both   p-values < desired level for I(1) variables)
```

23. Notice that Pesaran et al. (2001) calculate the AIC and BIC differently. Compared to Stata's computation, their statistics have the opposite sign and are divided by 2.

```
decision: no rejection (.a), inconclusive (.), or rejection (.r) at levels:

             |       10%         5%         1%
-------------+--------------------------------
    decision |         .          .          .

. quietly ardl w Prod UR Wedge Union if smpl, exog(D7475 D7579) lags(6 1 6 6 6)
> trendvar restricted ec

. estat ectest, norule

Pesaran, Shin, and Smith (2001) bounds test

H0: no level relationship                                        F =     4.780
Case 4                                                           t =    -2.437

Finite sample (4 variables, 104 observations, 26 short-run coefficients)

Kripfganz and Schneider (2020) critical values and approximate p-values

   | 10%              | 5%               | 1%               | p-value
   |    I(0)     I(1) |    I(0)     I(1) |    I(0)     I(1) |    I(0)     I(1)
---+------------------+------------------+------------------+-----------------
 F |   2.576    3.693 |   2.988    4.219 |   3.906    5.374 |   0.002    0.023
 t |  -2.954   -3.837 |  -3.281   -4.210 |  -3.922   -4.926 |   0.247    0.532

decision: no rejection (.a), inconclusive (.), or rejection (.r) at levels:

             |       10%         5%         1%
-------------+--------------------------------
    decision |        .a         .a         .a
```

At this stage, we are not yet interested in an interpretation of the coefficient estimates. In the top-right corner, the bounds test output displays the test statistics for the first two testing steps, as outlined in Section 2.4. Their values equal those reported by Pesaran et al. (2001, Table II).[24] While we are using the Kripfganz and Schneider (2020) critical values for finite samples instead of asymptotic critical values, the qualitative conclusions remain the same as in the original study. In the first step, we consider the F-statistic for the joint null hypothesis $\alpha = 0$ and $\sum_{i=0}^{p} \boldsymbol{\beta}_i = \mathbf{0}$. The test statistic is larger than the upper-bound critical values – which would be the exact critical values if all long-run forcing variables were $I(1)$ – for the 10% and 5% significance levels, both under case 3 and 4. If we take a more cautious stand with the 1% significance level, the result is inconclusive because the F-statistic falls within the two bounds. The easiest way to see this is by comparing the p-values in the final columns to the desired significance level.

Given that the evidence so far tends towards a rejection of the null hypothesis, we should consider the individual null hypothesis $\alpha = 0$ in the second step. Here, we notice substantial differences between the two cases. While the value of the t-statistic falls within the bounds for all three significance levels when we do not include a deterministic trend, we clearly cannot reject the null hypothesis when a trend is included. Overall, the statistical evidence is mixed. Given that the time trend itself is statistically insignificant, we could decide to go with the results for case 3. If we are lenient with the inconclusive result from the second step, we may conclude that there is mild support for the existence

24. Pesaran et al. (2001) additionally report the bounds test results for case 5 with an unrestricted trend, and for lag orders $p \in \{4, 5\}$. The respective test statistics can be replicated as well by adjusting the `ardl` command line accordingly.

of a long-run level relationship.

As a third step, we should check the statistical significance of the long-run coefficients $\boldsymbol{\theta}$. Instead of looking at the coefficient estimates from the above ARDL(6, 1, 6, 6, 6) model, Pesaran et al. (2001) aim for a more parsimonious specification with potentially different lag orders for each variable, using a maximum lag order $p^* = 6$ and the AIC as the model selection criterion. While the optimization over all 14,406 lag combinations finishes in virtually no time using our fast Mata algorithm, we illustrate how to display a progress bar with the option `dots`, which might be useful for larger models:

```
. ardl w Prod UR Wedge Union if smpl, exog(D7475 D7579) maxlag(6) aic ec1 dots

Optimal lag selection, % complete:
----+---20%---+---40%---+---60%---+---80%---+-100%
..................................................
AIC optimized over 14406 lag combinations

ARDL(6,0,5,4,5) regression

Sample: 1972q1 thru 1997q4                          Number of obs =    104
                                                    R-squared     = 0.6726
                                                    Adj R-squared = 0.5620
Log likelihood = 367.25283                          Root MSE      = 0.0082

------------------------------------------------------------------------------
         D.w | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
ADJ          |
           w |
         L1. |  -.2471572   .0521006    -4.74   0.000    -.3509028   -.1434115
-------------+----------------------------------------------------------------
LR           |
        Prod |
         L1. |   1.069227   .0451471    23.68   0.000     .9793278    1.159126
             |
          UR |
         L1. |  -.1010547   .0303894    -3.33   0.001    -.1615678   -.0405415
             |
       Wedge |
         L1. |  -.9322035   .2432158    -3.83   0.000    -1.416508   -.4478991
             |
       Union |
         L1. |   1.459416   .2847577     5.13   0.000     .8923912    2.026441
-------------+----------------------------------------------------------------
SR           |
           w |
         LD. |   -.418152   .0970869    -4.31   0.000    -.6114767   -.2248273
        L2D. |  -.3371241   .1076479    -3.13   0.002    -.5514784   -.1227698
        L3D. |  -.5355013   .1024435    -5.23   0.000    -.7394924   -.3315102
        L4D. |  -.1324765    .088904    -1.49   0.140     -.309507     .044554
        L5D. |  -.2017854   .0800474    -2.52   0.014    -.3611801   -.0423906
             |
        Prod |
         D1. |   .2642672   .0587165     4.50   0.000     .1473477    .3811866
             |
          UR |
         D1. |   .0038741    .008252     0.47   0.640    -.0125576    .0203059
         LD. |   .0180803   .0112588     1.61   0.112    -.0043388    .0404995
        L2D. |   .0064258   .0106366     0.60   0.548    -.0147544     .027606
        L3D. |   .0276765     .01116     2.48   0.015      .005454    .0498989
```

```
     L4D. |   .0304992   .0109952     2.77   0.007     .0086049    .0523935
          |
    Wedge |
      D1. |  -.3059897   .0515941    -5.93   0.000    -.4087266   -.2032527
      LD. |  -.0428407   .0584203    -0.73   0.466    -.1591703     .073489
     L2D. |  -.0922406   .0566866    -1.63   0.108    -.2051181    .0206369
     L3D. |  -.1886045   .0558292    -3.38   0.001    -.2997747   -.0774344
          |
    Union |
      D1. |  -.9557296   .8138693    -1.17   0.244     -2.57635    .6648913
      LD. |  -2.783417   .8141054    -3.42   0.001    -4.404508   -1.162326
     L2D. |  -.2560327   .8307343    -0.31   0.759    -1.910236    1.398171
     L3D. |   .0553435   .7432099     0.07   0.941    -1.424577    1.535264
     L4D. |  -2.185803   .6535687    -3.34   0.001    -3.487225   -.8843817
          |
    D7475 |   .0301089    .006154     4.89   0.000     .0178547     .042363
    D7579 |   .0169542   .0062481     2.71   0.008     .0045127    .0293958
    _cons |   .6604218   .1425601     4.63   0.000     .3765484    .9442952
------------------------------------------------------------------------------
```

The model with the smallest AIC value is an ARDL(6, 0, 5, 4, 5) model. The results are displayed in the EC representation (4). The first coefficient in the `ADJ` section is the negative speed-of-adjustment coefficient ($-\alpha$). The coefficients in the `LR` section are the long-run coefficients $\boldsymbol{\theta}$, and those in the `SR` section are the short-run coefficients $\boldsymbol{\omega}$, $\boldsymbol{\psi}_{xi}$, and $\boldsymbol{\gamma}$, together with the unrestricted intercept c_0.

The long-run coefficients are all highly statistically significant – our final check for the existence of a level relationship. For a detailed interpretation of the results, we refer the reader to Pesaran et al. (2001). We just note that it is not surprising that the long-run level effect of `Prod` on `w` is statistically not significantly different from unity. This is in line both with economic theory and the observed co-movement of the respective time series in Figure 2. The estimate of the speed-of-adjustment coefficient indicates that a disturbance to the long-run equilibrium is corrected by 25% within one quarter, which corresponds to a half life of about 2.4 quarters. Importantly, the reported p-value and confidence interval for the speed-of-adjustment coefficient should not be taken at face value. The t-statistic for this coefficient does not have a standard distribution under the null hypothesis that it equals zero. In fact, this is the test statistic which we considered under the second step of the bounds step. To obtain appropriate critical values and p-values (under a one-sided alternative hypothesis), we can simply call the `estat ectest` postestimation command again:

```
. estat ectest, norule

Pesaran, Shin, and Smith (2001) bounds test

H0: no level relationship                                    F =     7.367
Case 3                                                       t =    -4.744

Finite sample (4 variables, 104 observations, 21 short-run coefficients)

Kripfganz and Schneider (2020) critical values and approximate p-values

   | 10%               | 5%                | 1%                | p-value
   |    I(0)     I(1) |    I(0)     I(1) |    I(0)     I(1) |    I(0)     I(1)
---+-------------------+-------------------+-------------------+-----------------
 F |   2.392    3.638 |   2.838    4.212 |   3.835    5.469 |   0.000    0.001
 t |  -2.470   -3.532 |  -2.796   -3.901 |  -3.434   -4.605 |   0.000    0.007
```

```
decision: no rejection (.a), inconclusive (.), or rejection (.r) at levels:
             |       10%         5%         1%
-------------+---------------------------------
    decision |        .r         .r         .r
```

Since the p-value for the t-statistic is smaller than the conventional significance levels, the speed-of-adjustment estimate can be regarded as statistically significant. In contrast to the earlier ARDL(6, 1, 6, 6, 6) model, the bounds test from this ARDL(6, 0, 5, 4, 5) specification now provides unambiguous evidence in favor of a long-run level relationship.

The short-run coefficients reveal a picture of complicated dynamics. Evidently, a temporary change in most of the explanatory variables is not fully absorbed at once. A more detailed interpretation of individual coefficients does not seem very fruitful here. Instead, a forecasting exercise with the `forecast` command suite or a dynamic simulation of the effects with the community-contributed `dynardl` command (Jordan and Philips 2018) could provide useful insights. We provide a brief forecasting example at the end of this section. At this point, a technical comment is in order. As mentioned in Section 2.3, the lag order $q = 0$ for `Prod` introduced a perfect relationship among the effects in the latest regression output. Observe that the only short-run coefficient of `Prod` equals the product of its long-run coefficient with the speed-of-adjustment coefficient:

```
. nlcom -_b[ADJ:L.w] * _b[LR:L.Prod]
       _nl_1: -_b[ADJ:L.w] * _b[LR:L.Prod]
------------------------------------------------------------------------------
         D.w | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
       _nl_1 |   .2642672   .0587165     4.50   0.000      .149185    .3793493
------------------------------------------------------------------------------
```

This perfect link among the coefficients implies that there is no independent short-run response to a change in `Prod` after accounting for the equilibrium adjustment. The additional short-run coefficient is merely a consequence of formulating the long-run relationship in terms of the lagged regressors $\mathbf{x}_{t-1}$ – equation (4) – when there is no lag of `Prod` in the underlying ARDL model (3). To avoid this complication, we can re-parameterize the EC model as in equation (6). The long-run coefficients remain unchanged but the superfluous short-run effect disappears:

```
. ardl w Prod UR Wedge Union if smpl, exog(D7475 D7579) maxlag(6) aic ec
ARDL(6,0,5,4,5) regression
Sample: 1972q1 thru 1997q4                             Number of obs =    104
                                                       R-squared     = 0.6726
                                                       Adj R-squared = 0.5620
Log likelihood = 367.25283                             Root MSE      = 0.0082
------------------------------------------------------------------------------
         D.w | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
ADJ          |
           w |
         L1. |  -.2471572   .0521006    -4.74   0.000    -.3509028   -.1434115
-------------+----------------------------------------------------------------
```

```
LR           |
        Prod |   1.069227   .0451471    23.68   0.000     .9793278    1.159126
          UR |  -.1010547   .0303894    -3.33   0.001    -.1615678   -.0405415
       Wedge |  -.9322035   .2432158    -3.83   0.000    -1.416508   -.4478991
       Union |   1.459416   .2847577     5.13   0.000     .8923912    2.026441
-------------+----------------------------------------------------------------
SR           |
           w |
         LD. |   -.418152   .0970869    -4.31   0.000    -.6114767   -.2248273
        L2D. |  -.3371241   .1076479    -3.13   0.002    -.5514784   -.1227698
        L3D. |  -.5355013   .1024435    -5.23   0.000    -.7394924   -.3315102
        L4D. |  -.1324765    .088904    -1.49   0.140     -.309507     .044554
        L5D. |  -.2017854   .0800474    -2.52   0.014    -.3611801   -.0423906
             |
          UR |
         D1. |   .0288505    .009703     2.97   0.004     .0095294    .0481716
         LD. |   .0180803   .0112588     1.61   0.112    -.0043388    .0404995
        L2D. |   .0064258   .0106366     0.60   0.548    -.0147544     .027606
        L3D. |   .0276765     .01116     2.48   0.015      .005454    .0498989
        L4D. |   .0304992   .0109952     2.77   0.007     .0086049    .0523935
             |
       Wedge |
         D1. |  -.0755889    .057272    -1.32   0.191     -.189632    .0384543
         LD. |  -.0428407   .0584203    -0.73   0.466    -.1591703     .073489
        L2D. |  -.0922406   .0566866    -1.63   0.108    -.2051181    .0206369
        L3D. |  -.1886045   .0558292    -3.38   0.001    -.2997747   -.0774344
             |
       Union |
         D1. |  -1.316435   .7848436    -1.68   0.098    -2.879258    .2463886
         LD. |  -2.783417   .8141054    -3.42   0.001    -4.404508   -1.162326
        L2D. |  -.2560327   .8307343    -0.31   0.759    -1.910236    1.398171
        L3D. |   .0553435   .7432099     0.07   0.941    -1.424577    1.535264
        L4D. |  -2.185803   .6535687    -3.34   0.001    -3.487225   -.8843817
             |
       D7475 |   .0301089    .006154     4.89   0.000     .0178547     .042363
       D7579 |   .0169542   .0062481     2.71   0.008     .0045127    .0293958
       _cons |   .6604218   .1425601     4.63   0.000     .3765484    .9442952
------------------------------------------------------------------------------
```

Pesaran et al. (2001) initially constrained the lag orders to be identical in order to limit pretesting problems. The actual implications of the different lag selection approaches are unclear. As long as there is no irrefutable evidence of residual serial correlation, one may as well skip the testing based on equal lag orders. In our case, the serial-correlation test appears acceptable (although not too comfortably):

```
. estat bgodfrey, lags(1 4)
Breusch-Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |          2.463               1                   0.1166
       4     |          7.518               4                   0.1109
---------------------------------------------------------------------------
                        H0: no serial correlation
```

A closer look at the coefficient estimates from the ARDL(6, 0, 5, 4, 5) model reveals that they are very similar but not identical to those reported by Pesaran et al. (2001, Equation (31) and Table III). The reason for the discrepancy is that Pesaran et al. (2001) fixed the lag order for `Prod` at $q = 1$, and therefore estimated an ARDL(6, 1, 5, 4, 5) model, despite claiming otherwise. This can be easily verified by replicating their estimates with the following code:

```
. ardl w Prod UR Wedge Union if smpl, exog(D7475 D7579) maxlag(6) lags(. 1 . . .)
> aic ec1
  (output omitted)
```

This completes the narrow replication of the Pesaran et al. (2001) empirial application. A few critical remarks shall be added. It is not always transparent why the authors used their *researcher degrees of freedom* in the way they have done. They reserved 8 initial observations for the lag selection, but then initially set the maximum lag order only to $p^* = 7$. The model with $p = 8$ lags – identical for all variables besides `Prod` – actually attains an even lower AIC, with or without linear time trend. Qualitatively, the conclusions are largely unaffected. With $p = 8$, the bounds test is even more supportive of a long-run level relationship than the one based on a fixed lag order of $p = 6$:

```
. ardl w Prod UR Wedge Union if smpl, exog(D7475 D7579) lags(8 1 8 8 8) ec
  (output omitted)
. estat ectest, norule
Pesaran, Shin, and Smith (2001) bounds test
H0: no level relationship                                         F =     7.991
Case 3                                                            t =    -3.853
Finite sample (4 variables, 104 observations, 34 short-run coefficients)
Kripfganz and Schneider (2020) critical values and approximate p-values
   | 10%              | 5%               | 1%               | p-value
   |    I(0)     I(1) |    I(0)     I(1) |    I(0)     I(1) |    I(0)     I(1)
---+------------------+------------------+------------------+-----------------
 F |   2.315    3.659 |   2.755    4.249 |   3.744    5.556 |   0.000    0.000
 t |  -2.410   -3.447 |  -2.746   -3.830 |  -3.402   -4.564 |   0.003    0.048
decision: no rejection (.a), inconclusive (.), or rejection (.r) at levels:
             |       10%        5%         1%
-------------+--------------------------------
    decision |        .r        .r          .
```

The decision for a lower lag order might still be justified on the grounds of conserving *model degrees of freedom*, but then one could reduce the number of initial observations put aside. The choice of the maximum lag order is an important initial step, which should not be undervalued. For instance, without the requirement of equal lag orders, setting $p^* = 7$ or $p^* = 8$ yields an ARDL(4, 7, 5, 4, 5) or ARDL(7, 2, 5, 4, 8) model, respectively. The optimal lag orders for `UR` and `Wedge` are quite robust, while there is considerable variation for the remaining variables, which is a consequence of their very high persistence. This also indicates that restricting the lag order for `Prod` from the outset may not be justified.

After establishing the evidence in favor of a long-run level relationship, we may want to estimate a more parsimonious model by re-optimizing the lag order with the BIC:[25]

25. In the interest of conserving space, we do not show further regression output. Our example can be easily reconstructed with the publicly available data set.

```
. ardl w Prod UR Wedge Union, exog(D7475 D7579) maxlag(8) bic ec
  (output omitted)
. estat bgodfrey, lags(1 4)
Breusch{Godfrey LM test for autocorrelation
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |          0.664               1                   0.4151
       4     |          1.717               4                   0.7877
---------------------------------------------------------------------------
                        H0: no serial correlation
```

Even allowing for a maximum lag order of $p^* = 8$, the BIC selects an ARDL(4, 0, 0, 4, 2) model with a considerably smaller number of parameters than the model chosen by the AIC. Potential worries about serial error correlation due to omitted dynamics appear unfounded. The long-run coefficient of `Prod` remains statistically indistinguishable from unity. The other three long-run coefficients become smaller and they partially lose their statistical significance, although the quite wide confidence intervals for the long-run effects of `Wedge` and `Union` still cover a range of relevant effect sizes.

Further specification checks can be carried out using standard postestimation commands. For instance, it is common practice to check for coefficient stability with a cumulative sum (CUSUM) test:

```
. estat sbcusum, ylabel(, angle(horizontal))
> tlabel(1972q1(8)1997q4, format(%tqCCYY)) xtitle("") name(sb1)
Cumulative sum test for parameter stability
Sample: 1972q1 thru 1997q4                 Number of obs = 104
H0: No structural break
                   Test      -------- Critical value ---------
      Type    statistic           1%           5%          10%
--------------------------------------------------------------
 Recursive       0.8823       1.1430       0.9479       0.8499
--------------------------------------------------------------
. estat sbcusum, ols ylabel(, angle(horizontal))
> tlabel(1972q1(8)1997q4, format(%tqCCYY)) xtitle("") name(sb2)
Cumulative sum test for parameter stability
Sample: 1972q1 thru 1997q4                 Number of obs = 104
H0: No structural break
                   Test      -------- Critical value ---------
      Type    statistic           1%           5%          10%
--------------------------------------------------------------
 OLS             0.4331       1.6276       1.3581       1.2238
--------------------------------------------------------------
. graph combine sb1 sb2, ysize(2) xsize(5)
```

Neither the CUSUM test based on recursive residuals nor the one based on OLS residuals triggers a warning flag. If we had found evidence of a structural break, this would have cast doubt on the earlier bounds test results. If a specific break date can be identified, a remedy might be to restrict the attention to a subsample before or after the break, or to split the sample and carry out the test separately for both subsamples.

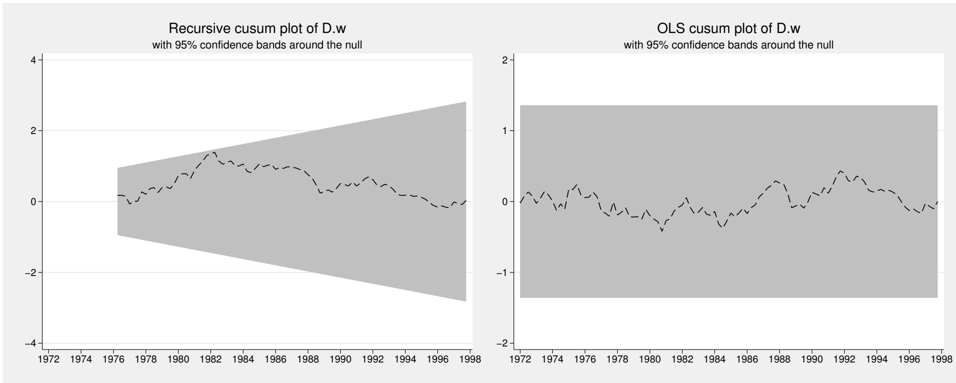

Figure 3: CUSUM plots

Other relevant tests could include checks for homoskedasticity and normality of the errors. While violations of these two conditions do not turn the OLS estimator inconsistent, they could invalidate the results from the bounds test. To keep it simple, we just look at the Cameron-Trivedi decomposition of the information matrix test:

```
. estat imtest
Cameron & Trivedi´s decomposition of IM-test
---------------------------------------------------
              Source |       chi2     df         p
---------------------+-----------------------------
  Heteroskedasticity |     104.00    103    0.4539
            Skewness |      29.46     16    0.0210
            Kurtosis |       0.00      1    0.9527
---------------------+-----------------------------
               Total |     133.46    120    0.1892
---------------------------------------------------
```

Concerns about heteroskedasticity seem unfounded. However, there is evidence of skewness, which indicates that the latest specification would not be appropriate for the bounds test. While Pesaran et al. (2001) have not done checks for normality, it turns out that choosing higher lag orders – as selected by the AIC – can alleviate this problem for our example. The bounds test results reported earlier thus remain reliable.

Finally, we conclude our example with an illustration of a forecasting exercise. A parsimonious model specification – as selected by the BIC – is appropriate for this task, because it tends to have better predictive performance in small samples. Let us say we are interested in the out-of-sample forecast of `w` for the next 8 quarters, holding fixed all exogenous variables at their last in-sample values. After appending the data set accordingly, we use the official `forecast` command suite to achieve this task:

```
. set obs `= _N + 8´
Number of observations (_N) was 116, now 124.
. replace qtr = qtr[_n-1] + 1 if qtr > tq(1997q4)
(8 real changes made)
```

```
. tsset qtr

Time variable: qtr, 1969q1 to 1998q4
Delta: 1 quarter

. foreach var of varlist Prod UR Wedge Union D7475 D7579 {
  2.         replace `var´ = L.`var´ if qtr > tq(1997q4)
  3. }
  (output omitted)

. quietly ardl w Prod UR Wedge Union, exog(D7475 D7579) maxlag(8) bic

. estimates store ardl

. forecast create ardl
  Forecast model ardl started.

. forecast estimates ardl, predict(xb)
  Added estimation results from ardl.
  Forecast model ardl now contains 1 endogenous variable.

. forecast exogenous Prod UR Wedge Union D7475 D7579
  Forecast model ardl now contains 6 declared exogenous variables.

. forecast solve, begin(tq(1998q1))

Computing dynamic forecasts for model ardl.
-------------------------------------------
Starting period:  1998q1
Ending period:    1999q4
Forecast prefix:  f_

1998q1:  ...........
1998q2:  ...........
1998q3:  ........
1998q4:  ...........
1999q1:  ...........
1999q2:  ...........
1999q3:  ..........
1999q4:  ...........

Forecast 1 variable spanning 8 periods.
---------------------------------------
```

In the following, we want to compare this baseline forecast with two alternative scenarios. In the first scenario, a positive shock of magnitude 0.01 hits `w` in the first out-of-sample quarter. In the second scenario, we consider a negative shock of the same magnitude.[26] Eventually, we plot the three forecasts in Figure 4:

```
. forecast adjust w = w + 0.01 if qtr == tq(1998q1)
  Endogenous variable w now has 1 adjustment.

. forecast solve, begin(tq(1998q1)) prefix(fp_)
  (output omitted)

. forecast adjust w = w - 0.02 if qtr == tq(1998q1)
  Endogenous variable w now has 2 adjustments.

. forecast solve, begin(tq(1998q1)) prefix(fn_)
  (output omitted)

. tsline w f_w fp_w fn_w if qtr >= tq(1997q1), ylabel(, angle(horizontal))
> tlabel(1997q1(2)1999q4) xtitle("")
```

The example can be easily adjusted to accommodate other scenarios, which might

26. Because the adjustment are made consecutively within the same forecast model, the second adjustment needs to be -0.02. Otherwise, we would only revert back to the baseline scenario.

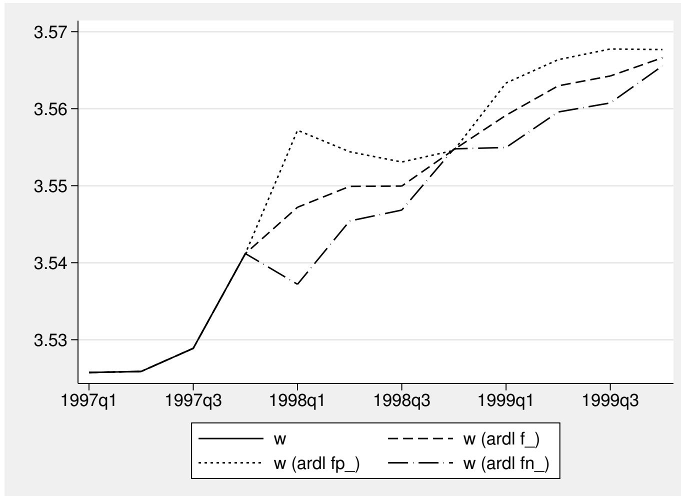

Figure 4: Time-series graph for `w` with different forecast scenarios

involve alternative time paths for the exogenous variables. We leave such exercises to the interested reader.

6 Conclusion

In this article, we have described the `ardl` Stata command for the estimation of autoregressive distributed lag models with time-series data. The lag orders can be prespecified or chosen optimally with the Akaike or Schwarz/Bayesian information criterion. For this purpose, the command is able to estimate tens of thousands of candidate models in virtually no time. Two useful reparameterizations of the model in error-correction form allow for an interpretation of the coefficients as short-run and long-run effects. The command further enables testing for the existence of a long-run level relationship using the popular bounds test, which is implemented as a postestimation feature. In the case of nonstationary variables, this amounts to cointegration testing. Yet, the ARDL approach is flexible to allow for both stationary and nonstationary variables. The package provides the recently improved Kripfganz and Schneider (2020) critical values for the bounds test, which allow accurate inference for almost all practically relevant combinations of sample size, number of long-run forcing variables, lag orders, and deterministic model components.

7 Acknowledgement

We thank Michael Binder for his support and guidance during early stages of this project. Moreover, we are grateful for numerous comments and suggestions from the

Stata community which helped to improve our `ardl` package. This includes countless e-mail communications, discussions on the *Statalist* forum, and exchanges of ideas at the 2016 Stata Conference in Chicago, the 2017 and 2018 German Stata user group meetings in Berlin and Konstanz, respectively, and the 2018 UK Stata Conference in London.

8 References

Adeleye, N., E. Osabuohien, E. Bowale, O. Matthew, and E. Oduntan. 2018. Financial reforms and credit growth in Nigeria: Empirical insights from ARDL and ECM techniques. *International Review of Applied Economics* 32(6): 807–820.

Ang, J. B. 2007. CO_2 emissions, energy consumption, and output in France. *Energy Policy* 35(10): 4772–4778.

———. 2010. Finance and inequality: The case of India. *Southern Economic Journal* 76(3): 738–761.

Bahmani-Oskooee, M., and T. J. Brooks. 1999. Bilateral J-curve between U.S. and her trading partners. *Weltwirtschaftliches Archiv* 135(1): 156–165.

Bekun, F. V., F. Emir, and S. A. Sarkodie. 2019. Another look at the relationship between energy consumption, carbon dioxide emissions, and economic growth in South Africa. *Science of the Total Environment* 655: 759–765.

Belloumi, M. 2014. The relationship between trade, FDI and economic growth in Tunisia: An application of the autoregressive distributed lag model. *Economic Systems* 38(2): 269–287.

Blackburne, E. F., and M. W. Frank. 2007. Estimation of nonstationary heterogeneous panels. *Stata Journal* 7(2): 197–208.

Büyükşahin, B., and M. A. Robe. 2014. Speculators, commodities and cross-market linkages. *Journal of International Money and Finance* 42: 38–70.

Ciaian, P., M. Rajcaniova, and d'A. Kancs. 2016. The economics of BitCoin price formation. *Applied Economics* 48: 1799–1815.

———. 2018. Virtual relationships: Short- and long-run evidence from BitCoin and altcoin markets. *Journal of International Financial Markets, Institutions and Money* 52: 173–195.

Davis, E. P., and H. Zhu. 2011. Bank lending and commercial property cycles: Some cross-country evidence. *Journal of International Money and Finance* 30(1): 1–21.

De Vita, G., and A. Abbott. 2004. Real exchange rate volatility and US exports: An ARDL bounds testing approach. *Economic Issues* 9(1): 69–78.

Dickey, D. A., and W. A. Fuller. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74(366): 427–431.

Ditzen, J. 2018. Estimating dynamic common-correlated effects in Stata. *Stata Journal* 18(3): 585–617.

———. 2021. Estimating long-run effects and the exponent of cross-sectional dependence: An update to xtdcce2. *Stata Journal* 21(3): 687–707.

Engle, R. F., and C. W. J. Granger. 1987. Co-integration and error correction: representation, estimation, and testing. *Econometrica* 55(2): 251–276.

Enisan, A. A., and A. O. Olufisayo. 2009. Stock market development and economic growth: Evidence from seven sub-Sahara African countries. *Journal of Economics and Business* 61(2): 162–171.

Fatai, K., L. Oxley, and F. G. Scrimgeour. 2004. Modelling the causal relationship between energy consumption and GDP in New Zealand, Australia, India, Indonesia, The Philippines and Thailand. *Mathematics and Computers in Simulation* 64(3–4): 431–445.

Fedderke, J. W., P. Perkins, and J. M. Luiz. 2006. Infrastructural investment in long-run economic growth: South Africa 1875–2001. *World Development* 34(6): 1037–1059.

Frankel, J., S. L. Schmukler, and L. Servén. 2004. Global transmission of interest rates: monetary independence and currency regime. *Journal of International Money and Finance* 23(5): 701–733.

Halicioglu, F. 2009. An econometric study of CO_2 emissions, energy consumption, income and foreign trade in Turkey. *Energy Policy* 37(3): 1156–1164.

Hassler, U., and J. Wolters. 2006. Autoregressive distributed lag models and cointegration. *Allgemeines Statistisches Archiv* 90(1): 59–74.

Jalil, A., and S. F. Mahmud. 2009. Environment Kuznets curve for CO_2 emissions: A cointegration analysis for China. *Energy Policy* 37(12): 5167–5172.

Jordan, S., and A. Q. Philips. 2018. Cointegration testing and dynamic simulations of autoregressive distributed lag models. *Stata Journal* 18(4): 902–923.

Katircioglu, S. T. 2009. Revisiting the tourism-led-growth hypothesis for Turkey using the bounds test and Johansen approach for cointegration. *Tourism Management* 30(1): 17–20.

Kripfganz, S., and V. Sarafidis. 2021. Instrumental-variable estimation of large-T panel-data models with common factors. *Stata Journal* 21(3): 659–686.

Kripfganz, S., and D. C. Schneider. 2020. Response surface regressions for critical value bounds and approximate p-values in equilibrium correction models. *Oxford Bulletin of Economics and Statistics* 82(6): 1456–1481.

Morley, B. 2006. Causality between economic growth and immigration: An ARDL bounds testing approach. *Economics Letters* 90(1): 72–76.

Narayan, P. K. 2005. The saving and investment nexus for China: evidence from cointegration tests. *Applied Economics* 37(17): 1979–1990.

Narayan, P. K., and R. Smyth. 2005. Electricity consumption, employment and real income in Australia evidence from multivariate Granger causality tests. *Energy Policy* 33(9): 1109–1116.

Narayan, P. K., R. Smyth, and M. Nandha. 2004. Interdependence and dynamic linkages between the emerging stock markets of South Asia. *Accounting & Finance* 44(3): 419–439.

Ntanos, S., M. Skordoulis, G. Kyriakopoulos, G. Arabatzis, M. Chalikias, S. Galatsidas, A. Batzios, and A. Katsarou. 2018. Renewable energy and economic growth: Evidence from European countries. *Sustainability* 10(8): 2626.

Obstfeld, M., J. C. Shambaugh, and A. M. Taylor. 2005. The trilemma in history: Tradeoffs among exchange rates, monetary policies, and capital mobility. *Review of Economics and Statistics* 87(3): 423–438.

Oteng-Abayie, E. F., and J. M. Frimpong. 2006. Bounds testing approach to cointegration: An examination of foreign direct investment trade and growth relationships. *American Journal of Applied Sciences* 3(11): 2079–2085.

Ozturk, I. 2010. A literature survey on energy-growth nexus. *Energy Policy* 38(1): 340–349.

Payne, J. E. 2010. Survey of the international evidence on the causal relationship between energy consumption and growth. *Journal of Economic Studies* 37(1): 53–95.

Persyn, D., and J. Westerlund. 2008. Error-correction-based cointegration tests for panel data. *Stata Journal* 8(2): 232–241.

Pesaran, M. H., and Y. Shin. 1998. An autoregressive distributed-lag modelling approach to cointegration analysis. In *Econometrics and Economic Theory in the 20th Century. The Ragnar Frisch Centennial Symposium*, ed. S. Strøm, chap. 11, 371–413. Cambridge: Cambridge University Press.

Pesaran, M. H., Y. Shin, and R. J. Smith. 2001. Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics* 16(3): 289–326.

Phillips, P. C. B., and B. E. Hansen. 1990. Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* 57(1): 99–125.

Sari, R., S. Hammoudeh, and U. Soytas. 2010. Dynamics of oil price, precious metal prices, and exchange rate. *Energy Economics* 32(2): 351–362.

Schaffer, M. E. 2010. egranger: Stata module to perform Engle-Granger cointegration tests and 2-step ECM estimation. Statistical Software Components S457210, Boston College.

Shambaugh, J. C. 2004. The effect of fixed exchange rates on monetary policy. *Quarterly Journal of Economics* 119(1): 301–352.

Shin, Y., B. Yu, and M. Greenwood-Nimmo. 2014. Modelling asymmetric cointegration and dynamic multipliers in a nonlinear ARDL framework. In *Festschrift in Honor of Peter Schmidt: Econometric Methods and Applications*, ed. W. C. Horrace and R. C. Sickles, chap. 9, 281–314. New York: Springer.

Song, H., S. Lin, S. F. Witt, and X. Zhang. 2011. Impact of financial/economic crisis on demand for hotel rooms in Hong Kong. *Tourism Management* 32(1): 172–186.

Wang, Q. 2012. Long-run covariance and its applications in cointegration regression. *Stata Journal* 12(3): 515–542.

Wang, Y.-S. 2009. The impact of crisis events and macroeconomic activity on Taiwan's international inbound tourism demand. *Tourism Management* 30(1): 75–82.

Wolde-Rufael, Y. 2005. Energy demand and economic growth: The African experience. *Journal of Policy Modeling* 27(8): 891–903.

———. 2006. Electricity consumption and economic growth: A time series experience for 17 African countries. *Energy Policy* 34(10): 1106–1114.

Zhang, X., E. A. Davidson, D. L. Mauzerall, T. D. Searchinger, P. Dumas, and Y. Shen. 2015. Managing nitrogen for sustainable development. *Nature* 528: 51–59.

About the authors

Sebastian Kripfganz joined the University of Exeter Business School in September 2015 as a Lecturer in Economics after finishing his Ph.D. at Goethe University Frankfurt. He has a strong research interest in methods for estimation and inference in dynamic econometric models.

Daniel Schneider received his Ph.D. in Economics from Goethe University Frankfurt in 2013 and is now a Statistical Programmer and Research Scientist at the Max Planck Institute for Demographic Research in Rostock. His research interests include time-series analysis and discrete-time multi-state models.