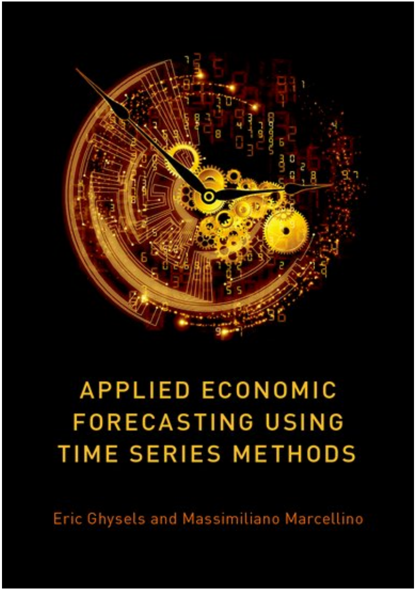
APPLIED ECONOMIC
FORECASTING USING
TIME SERIES METHODS
Eric Ghysels and Massimiliano Marcellino

# Applied Economic Forecasting using Time Series Methods

**Eric Ghysels and Massimiliano Marcellino**

Companion Slides - Chapter 3 The Dynamic Linear Regression Model

# The Baseline Linear Regression Model - Outline

# Overview

- We now turn our attention to models with dynamics, that is models with multiply time periods in an equation.
- First, we outline several popular dynamic linear regression models.
- We then consider issues surrounding model estimation and specification.
- Finally, examples with both simulated and empirical data are considered.

# ARDL examples

- The most general specification for a dynamic linear model is the auto regressive distributed lags model (ARDL) in which $y$, the dependent variable, is allowed to depend on $p$ lags of itself $(y_{t-1}, ..., y_{t-p})$ and $q$ lags of the regressors $x_t$ $(x_{t-1}, ..., x_{t-q})$. The simplest specification is the ARDL(1,1)

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \alpha_1 y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} (0, \sigma_\varepsilon^2). \tag{1}$$

- Several commonly used model originate from (1) with some added restrictions. We consider several examples.

# ARDL examples con't

- **Static Regression** - If we further assume $\alpha_1 = \beta_2 = 0$, then we see that ARDL nests the static linear regression discussed in the previous chapter.

$$\textit{Static regression: } \alpha_1 = \beta_2 = 0 \Longrightarrow y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \qquad (2)$$

  - Examples of static regression include the Keynesian consumption model, IS-LM, PPP, and some no arbitrage conditions for financial markets.
  - Mainly, static regressions are useful for modeling long-run relationships.

# ARDL examples con't

- **Autoregressive model of order one ($AR(1)$)** - Simplest time series model and shows up throughout the course.

$$\textit{AR(1) model}: \beta_1 = \beta_2 = 0 \Rightarrow y_t = \beta_0 + \alpha_1 y_{t-1} + \varepsilon_t \tag{3}$$

- The **Random walk model** is a specific case of the $AR(1)$ model in which $\alpha_1 = 1$.

$$\textit{Random walk model:}\ \beta_i = 0, i = 0, 1, 2\ \alpha_1 = 1 \Rightarrow y_t = y_{t-1} + \varepsilon_t \tag{4}$$

  - Restricting $\alpha_1$ to $1$ completely changes the dynamics of the AR process.
  - Efficient capital market theories generate an example of a random walk.

# ARDL examples con't

- The **differenced model** allows us to view how a change in $y_t$, $\Delta y_t = y_t - y_{t-1}$, is explained by a change in $x_t$, $\Delta x_t = x_t - x_{t-1}$

$$\textit{Differenced model}: \alpha_1 = 1,\ \beta_1 = -\beta_2 \Rightarrow \Delta y_t = \beta_0 + \beta_1 \Delta x_t + \varepsilon_t \quad (5)$$

  - Differenced models can be use to net out the effects of trends that some economic time series display.
  - Looking at differences in the log gives us an approximation for the growth rate of a variable (e.g. linking the growth rate of consumption to the growth rate of disposable income).

# ARDL examples con't

- The **leading indicator model** assumes x is a leading indicator for y.

$$\textit{Leading indicator model}: \alpha_1 = \beta_1 = 0 \Rightarrow y_t = \beta_0 + \beta_2 x_{t-1} + \varepsilon_t \quad (6)$$

  - e.g. consumers confidence could be a leading indicator for GDP growth rates.

- The **distributed lag** $(DL)$ model can capture the effects of adjustment costs and other types of friction that have a dynamic structure and do not impact only instantaneously on y.

$$\textit{Distributed lag model:} \ \alpha_1 = 0 \Rightarrow y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + \varepsilon_t \quad (7)$$

# ARDL examples con't

- We can generalize the DL model discribed by (7) by adding lags of the independent variable. For example, the **geometric distributed lags** $(GDL)$ model has the following specification:

$$\begin{aligned} y_t &= \alpha + \beta(x_t + \omega x_{t-1} + \omega^2 x_{t-2+\ldots}) + \varepsilon_t \qquad 0 < \omega < 1, \\ &= \alpha + \beta \sum_{i=0}^{\infty} \omega^i x_{t-i} + \varepsilon_t. \end{aligned} \tag{8}$$

  - The long run effect of a change in $x$ on $y$ is $\beta \sum_0^\infty \omega^i = \beta/(1-\omega)$.
  - The closer $\omega$ is to one, the greater the difference is between the long run multiplier and the impact multiplier, $\beta$.

# ARDL examples con't

- The GDL model in (8) can be rewritten as:

$$y_t - \omega y_{t-1} = \alpha(1-\omega) + \beta x_t + \varepsilon_t \quad \varepsilon_t = \varepsilon_t - \omega\varepsilon_{t-1} \tag{9}$$

  - Note, it we set $\beta = \beta_0$ and $\omega = \alpha_1$ then (9) becomes a special case of model ARDL(1,1) in equation (1).

# ARDL examples con't

- The **partial adjustment model** model takes its name from the procedure which enable us to derive it.

*Partial adjustment model:* $\beta_2 = 0 \Rightarrow y_t = \beta_0 + \beta_1 x_t + \alpha_1 y_{t-1} + \varepsilon_t$ (10)

  - Suppose that $y$ has a target level, indicated by $y^*$ :

$$y_t^* = \alpha' + \beta' x_t + \varepsilon_t', \tag{11}$$

  - We assume that $y$ cannot be instantaneously modified to reach the target $y_t^*$ due to technical reasons or transaction costs. We then have

$$y_t - y_{t-1} = \gamma(y_t^* - y_{t-1}) \qquad 0 < \gamma < 1. \tag{12}$$

- The closer $\gamma$ is to $1$, the faster the adjustment of $y_t$ to $y_t^*$. Moreover, if we substitute (11) in (12) we obtain

$$y_t = \underbrace{\gamma\alpha'}_{\beta_0} + \underbrace{\gamma\beta'}_{\beta_1} x_t + \underbrace{(1-\gamma)}_{\alpha_1} y_{t-1} + \underbrace{\gamma\varepsilon'}_{\varepsilon_t}, \tag{13}$$

which is a specification similar to (9).

# ARDL examples con't

- The Error Correction Model (ECM) is a widely used specification in modern econometrics.

$$\textit{Error correction model} \quad \beta_1 + \beta_2 + \alpha_1 = 1 \Rightarrow$$
$$\Delta y_t = \beta_0 + \beta_1 \Delta x_t + (1 - \alpha_1)(y_{t-1} - x_{t-1}) + \varepsilon_t \qquad (14)$$

  - The intuition behind it is that in the long-run $y$ and $x$ are moving together so that deviations of $y$ from $x$ cannot persist over time and will gradually be evaporated.
  - As a consequence, $y$, and possibly also $x$, change not only due to changes in the exogenous variable $x$ but also as a consequence of changes in the deviations from the long-run equilibrium $y - x$.

- In the **dead start model** the independt variable $x$ has no contemporaneous effect on $y$.

$$\textit{Dead start model}: \beta_1 = 0 \Rightarrow y_t = \beta_0 + \beta_2 x_{t-1} + \alpha_1 y_{t-1} + \varepsilon_t \qquad (15)$$

  - If we model the dynamic of $x$ as

$$x_t = \gamma_0 + \gamma_1 x_{t-1} + \alpha_2 y_{t-1} + v_t, \qquad (16)$$

  then the two equations (15) and (16) considered jointly give rise to the so called Vector Autoregressive (VAR) model, which is the multivariate equivalent of the autoregressive model (3).

# ARDL examples con't

- The **autoregressive errors model** is a special case of an AD(1,1) model with a non-linear restriction on the parameters, which is commonly know as a Common Factor Restriction (COMFAC), namely $\alpha_1\beta_1 + \beta_2 = 0$.

$$y_t = \gamma_0 + \beta_1 x_t + u_t, u_t = \alpha_1 u_{t-1} + \varepsilon_t \tag{17}$$

- One of the requirements for the OLS estimator covered in Chapter 1 was that independent variables and the error terms be independent or at least uncorrelated. This assumption breaks down in dynamic linear models.

- Consider the $AR(1)$ model:

$$y_t = \alpha_1 y_{t-1} + \varepsilon_t. \tag{18}$$

By substituting for $y_{t-1}$ we obtain:

$$y_t = \alpha_1^2 y_{t-2} + \alpha_1 \varepsilon_{t-1} + \varepsilon_t.$$

By iterating this backward substitution we have:

$$y_t = \varepsilon_t + \alpha_1 \varepsilon_{t-1} + \alpha_1^2 \varepsilon_{t-2} + \alpha_1^3 \varepsilon_{t-3} + \ldots$$

By defining $x_t = (y_{t-1}, y_{t-2}, y_{t-3}, \ldots)'$ and $e_t = (\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots)'$, it is evident that $x_t$ and $e_t$ are not independent. Hence, in general, the OLS estimator will be biased.

# Estimating and Testing

- For example, let us consider the AR(1) model (18) and assume that the errors evolve according to:

$$\varepsilon_t = e_t - \omega e_{t-1}. \tag{19}$$

where $e_t$ is $\overset{iid}{\sim} (0, \sigma_e^2)$. In this case $y_{t-1}$ is correlated with $\varepsilon_{t-1}$, hence we need to find some valid instruments for $y_{t-1}$ and use an instrumental variable (IV) estimator.

# Estimating and Testing

- After parameter estimators are available, inference can be carried out by means of $t$ or $F$ statistics.
- In general, it is better to rely on the asymptotic distribution of these statistics rather than the finite sample properties.
- For example, for testing $q$ hypotheses on the model parameters, instead of using an $F-$statistic we should use $qF \underset{H_0}{\overset{a}{\sim}} \chi^2(q)$.

# General to Specific Specification Search

- Model specification is similar to the problem of variable selection in the context of the linear regression model.
- One approach for model specification called a *general to specific specification search* involves hypothesis tests sequentially applied to verify whether subsequent lags are significant.
- For example, starting with an ARDL(p,p) model, it is possible to test whether $y_{t-p}$ and $x_{t-p}$ are both insignificant.
- If the null is not rejected, then one considers an ARDL(p-1,p-1) model and tests the significance of $y_{t-p+1}$ and $x_{t-p+1}$. The tests stop when one finally rejects the null.

# General to Specific Specification Search

- What should be used as the initial lag length $p$?
  - If $p$ is too high w.r.t. the “true” data generating process, there will be multicollinearity problems among the regressors and loss of efficiency in the estimates.
  - If $p$ is lower than in the “true” data generating process, the estimates will be inconsistent.
- One criterion to select $p$ could depend on the data frequency (e.g, $p = 4$ for quarterly data) combined with a check for no mis-specification of the initial ARDL(p,p) model.
- However, iterative processes should be used with caution, being a statistical short-cut rather than a rigorously formalized procedure.
- LASSO and LARS could also be used in this context.

# Information Criteria

- **Information criteria** combine a measure of the model goodness of fit with a penalty accounting for the number of parameters.
  - A general form for an information criterion is

$$\log(\sigma^2) + g(k, T) \tag{20}$$

    where $\sigma^2$ is (an estimator of) the model error variance and $g(k, T)$ is a function of the number of parameters ($k$) and observations ($T$).

# Information Criteria

- The most widely used information criteria are:

$$\begin{array}{rl} \text{Akaike information criterion (AIC):} & g(k,T) = 2k/T \\ \text{Schwarz information criterion (BIC):} & g(k,T) = k\log(T)/T \\ \text{Hannan-Quinn information criterion (HQ):} & g(k,T) = 2k\log\log(T)/T \end{array}$$

  - BIC and HQ select the model with probability approaching one as T goes to infinity.
  - However, it is not possible to provide a uniformly valid ranking for the appropriateness of these criteria in finite samples, so that in applications a comparison of alternative IC can be useful.

# Other Model Specification Issues

- The hypotheses underlying the model should also hold within the dynamic linear regression model framework.
- Diagnostics on error terms to check for no autocorrelation, homoskedasticity, linearity, and normality are still applicable in this more general context.
- Parameter stability tests and the formulation of the dummy variables also need to be properly modified to account for the dynamic structure.
  - e.g in the AR(1) model an impluse dummy variable at time period $t$ will have an impact not only on $y_t$, but also on all subsequent periods.

# The Baseline Linear Regression Model - Outline

# Forecasting with Dynamic Models

- Let us illustrate forecasting with dynamic models using the ARDL(1,1) specification:

$$y_t = \beta_0 x_t + \beta_1 x_{t-1} + \alpha_1 y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} (0, \sigma_\varepsilon^2).$$

- To start with, we make the additional assumption that $\alpha_1 = 0$, so that the model simplifies to ARDL(0,1):

$$y_t = \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t.$$

- If we now group $x_t$ and $x_{t-1}$ into $z_t = (x_t, x_{t-1})$ and the parameters into $\beta = (\beta_0, \beta_1)'$, we can rewrite the model as

$$y_t = z_t \beta + \varepsilon_t.$$

- We can then use the exact same reasoning as in the linear regression model to show that the optimal (in the MSFE sense) forecast for $y_{T+h}$ is
$$\hat{y}_{T+h} = z_{T+h}\hat{\beta}, \tag{21}$$
where $\hat{\beta}$ is the OLS estimator of $\beta$.
- If the future values $z_{T+h}$ are unknown, they too should be replaced by forecasts.

# Forecasting with Dynamic Models

- Let us now add $y_{t-1}$ back, and write the ARDL(1,1) model as

$$y_t = z_t\beta + \alpha_1 y_{t-1} + \varepsilon_t.$$

- The optimal one-step ahead forecast is simply

$$\hat{y}_{T+1} = z_{T+1}\hat{\beta} + \widehat{\alpha}_1 y_T.$$

- For the two-steps ahead forecast, we would have

$$\hat{y}_{T+2} = z_{T+2}\hat{\beta} + \widehat{\alpha}_1 y_{T+1},$$

  but $y_{T+1}$ is unknown.

- Therefore, we replace it with $\hat{y}_{T+1}$, its conditional expectation at time $T$. The optimal forecast for $y_{T+2}$ is

$$\hat{y}_{T+2} = z_{T+2}\hat{\beta} + \widehat{\alpha}_1 \hat{y}_{T+1}.$$

- The optimal h-steps ahead forecast is

$$\hat{y}_{T+h} = z_{T+h}\hat{\beta} + \widehat{\alpha}_1 \hat{y}_{T+h-1}. \tag{22}$$

# Forecasting with Dynamic Models

- The forecast error is

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h}. \tag{23}$$

- Assuming that both $z_{T+h}$ and the parameters are known, so that there is not estimation uncertainty, we can write

$$e_{T+h} = \alpha_1(y_{T+h-1} - \hat{y}_{T+h-1}) + \varepsilon_{T+h} = \alpha_1 e_{T+h-1} + \varepsilon_{T+h}. \tag{24}$$

- Therefore, the presence of an autoregressive component in the model creates correlation in the h-steps ahead forecast error.

- If in (24) we replace $e_{T+h-1}$ with its expression we obtain

$$e_{T+h} = \alpha_1^2 e_{T+h-2} + \alpha_1 \varepsilon_{T+h-1} + \varepsilon_{T+h}.$$

- Repeated substitution then yields

$$e_{T+h} = \alpha_1^{h-1} \varepsilon_{T+1} + ... + \alpha_1 \varepsilon_{T+h-1} + \varepsilon_{T+h}, \tag{25}$$

since $e_{T+h} = 0$ for $h \leq 0$, which is an alternative representation for the dependence in $e_{T+h}$.

# Forecasting with Dynamic Models

- From (25) we can easily derive that

$$\begin{aligned} E(e_{T+h}) &= 0, \\ \text{Var}(e_{T+h}) &= (1 + \alpha_1^2 + ... + \alpha_1^{2(h-1)})\sigma_\varepsilon^2. \end{aligned}$$

- If the parameters are unknown, as well as future values of the $x$, then the forecast error becomes

$$\begin{aligned} e_{T+h} &= (z_{T+h}\beta - \widehat{z}_{T+h}\hat{\beta}) + (\alpha_1 y_{T+h-1} - \widehat{\alpha}_1 \hat{y}_{T+h-1}) + \varepsilon_{T+h} \\ &= z_{T+h}(\beta - \hat{\beta}) + (z_{T+h} - \widehat{z}_{T+h})\hat{\beta} + (\alpha_1 - \widehat{\alpha}_1)y_{T+h-1} \\ &\quad + \widehat{\alpha}_1 e_{T+h-1} + \varepsilon_{T+h} \end{aligned} \tag{26}$$

- Add last paragraph here

# Examples with simulated data

- We consider the same simulated data used the first two chapters. Recall that tests for no serial correlation rejected the null hypothesis that the model is static.
- Therefore, we now add lagged variables to the set of reressors in search for the best dynamic model specification, and build forecasts based on it.

# Examples with simulated data

- We focus on the following set of equations:

$$\begin{aligned} y_t \quad &= \alpha_1 + \beta_1 D_t + \alpha_2 x_t + \beta_2 D_t x_t + \gamma_1^y y_{t-1} \\ &+ \ldots + \gamma_p^y y_{t-p} + \gamma_1^x x_{t-1} + ... + \gamma_m^x x_{t-m} + \varepsilon_t \end{aligned}$$

  where $p = 0, 1, 2$ is the number of lags for $y$, and $m = 0, 1, 2$ is the number of lags of the independent variable $x$, so that the most general model for $y$ is an ARDL(2,2).
- We also include the dummy variable, $D_t$, which is equal to 0 for observations from 101 to 201 and 1 otherwise. According to the analysis in Chapter 2 the dummy variable is significant.

# Examples with simulated data

- We estimate all the models and then select the specification with the lowest Akaike Information Criterion (AIC) and Schwarz Information Criterion (SC).
- The table sums up the information criteria for all the model specifications tested:

| Model | AIC | Schwarz |
|---|---|---|
| DGP22 | 2.968 | 3.100 |
| DGP12 | 2.959 | 3.075 |
| DGP21 | 2.967 | 3.083 |
| DGP11 | 2.958 | 3.057 |
| DGP10 | 4.536 | 4.619 |
| DGP01 | 5.906 | 5.989 |
| DGP20 | 3.794 | 3.893 |
| DGP02 | 4.950 | 5.049 |

Table: *Dynamic model: Model selection*

# Examples with simulated data

- According to Table 1, the ARDL(1,1) is selected by both information criteria.
- This model actually coincides with our DGP:

$$y = \alpha_1 + \beta_1 D_t + \alpha_2 x_t + \beta_2 D_t x_t + \gamma_1^y y_{t-1} + \gamma_1^x x_{t-1} + \varepsilon_t$$

where $y = \ln(Y)$ from Chapter 1.

# Examples with simulated data

- Table 2 (next slide) presents the estimation output. All the coefficients are statistically significant (at the 1% confidence level).
- We can not reject the hypothesis that each coefficient is equal to its actual value in the DGP, and the $R^2$, is equal to $99.9\%$, while the Durbin-Watson statistic indicates that there is no serial correlation in the residuals.

# Estimation output

| | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| ALPHA(1) | 1.030 | 0.109 | 9.408 | 0.000 |
| BETA(1) | 0.920 | 0.155 | 5.954 | 0.000 |
| ALPHA(2) | 0.969 | 0.018 | 54.381 | 0.000 |
| BETA(2) | 1.005 | 0.026 | 38.020 | 0.000 |
| GAMMA-Y(1) | 0.496 | 0.008 | 59.533 | 0.000 |
| GAMMA-X(1) | 0.494 | 0.018 | 27.499 | 0.000 |
| | | | | |
| R-squared | 0.993 | Mean dep var | | 6.695 |
| Adjusted R-squared | 0.993 | S.D. dep var | | 12.560 |
| S.E. of regression | 1.046 | Akaike IC | | 2.958 |
| Sum squared resid | 212.418 | Schwarz IC | | 3.057 |
| Log likelihood | -289.812 | Hannan-Quinn | | 2.998 |
| F-statistic | 5695.113 | DW stat | | 2.219 |
| Prob(F-statistic) | 0.000 | | | |

Table: *DGP model estimation*

# Examples with simulated data: forecasting

- The final step of the analysis is forecasting. We take 301-501 as the forecast sample.
- The dynamic model forecasts for periods T+1 until T+H, where $h$=200.
- The static forecasts are a set of one-step ahead forecasts. while the latter are forecasts for periods T+1 until T+h, where $h$ =200.
- We also consider one-step ahead recursive forecasts, for the same forecast period.

# Examples with simulated data: forecasting

- Table 3 compares the forecasting performance indicators for the two types of forecasts.
- In comparison to the forecast results in Chapter 2 we see a vast improvement. Also noteworthy is the fact that the S.E. of the regression reported in Table 2 is very close to the RMSFE reported in Table 3 – a consequence of the fact that the parameter estimates are extremely precise.

| | Forecasting Method | |
|---|---|---|
| | *Static* | *Recursive* |
| *RMSFE* | 1.109 | 0.966 |
| *MAFE* | 0.865 | 0.761 |

Table: *One-step ahead ARDL(1,1) forecasting performance*

- The empirical examples using Euro-area GDP we considered in the previous chapters are based on simple linear regression without any lagged dependent or independent variables.
- If lagged variables have significant explanatory power but are excluded from the model, a likely impact will be that the omitted dynamics will appear in the errors that will be serially correlated as a result.
- The tests ran in the previous chapter rejected the null of serially uncorrelated errors. Inclusion of lagged variables helps capture the dynamics.

# Forecasting Euro area GDP growth

- The first model we consider, referred to as ARDL Model 1, includes contemporaneous values and the first lags of all variable:
  - Quarterly EA17 GDP growth $(y_t)$
  - Quarterly EA17 IP growth $(ipr_t)$
  - Quarterly growth rate of the Eurostoxx $(sr_t)$
  - Quarterly growth rate of the EA ESI $(su_t)$
- Not all lagged terms included in the ARDL Model 1 are significant.
  - The lagged IP growth does not appear to have much explanatory power.
  - The contemporaneous change of ESI also does not appear to be significant.

# ARDL Model 1 estimation

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 0.180 | 0.046 | 3.922 | 0.000 |
| Y(-1) | 0.287 | 0.116 | 2.473 | 0.016 |
| IPR | 0.261 | 0.034 | 7.696 | 0.000 |
| IPR(-1) | -0.012 | 0.049 | -0.250 | 0.803 |
| SU | -0.009 | 0.017 | -0.539 | 0.592 |
| SU(-1) | -0.019 | 0.017 | -1.115 | 0.269 |
| SR | 0.006 | 0.005 | 1.238 | 0.220 |
| SR(-1) | 0.007 | 0.005 | 1.384 | 0.171 |
| R-squared | 0.838 | Mean dep var | | 0.350 |
| Adjusted R-squared | 0.820 | S.D. dep var | | 0.629 |
| S.E. of regression | 0.267 | Akaike IC | | 0.304 |
| Sum squared resid | 4.422 | Schwarz IC | | 0.561 |
| Log likelihood | -2.656 | Hannan-Quinn | | 0.407 |
| F-statistic | 45.809 | DW stat | | 2.138 |
| Prob(F-statistic) | 0.000 | | | |

Table: *ARDL Model 1*

# Forecasting Euro area GDP growth

- We tried out a few alternative specifications based on reducing the ARDL Model 1, using information criteria to select the appropriate model.
- The outcome is ARDL Model 2, presented in Table 11, which includes the contemporaneous IP growth, the first lag of GDP growth, first lag of both the growth of ESI, and stock returns.

# ARDL Model 2 estimation

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 0.187 | 0.039 | 4.747 | 0.000 |
| Y(-1) | 0.273 | 0.065 | 4.174 | 0.000 |
| IPR | 0.262 | 0.027 | 9.827 | 0.000 |
| SU(-1) | -0.025 | 0.012 | -2.000 | 0.050 |
| SR(-1) | 0.008 | 0.005 | 1.673 | 0.099 |
| R-squared | 0.834 | Mean dep var | | 0.350 |
| Adjusted R-squared | 0.823 | S.D. dep var | | 0.629 |
| S.E. of regression | 0.264 | Akaike IC | | 0.245 |
| Sum squared resid | 4.538 | Schwarz IC | | 0.405 |
| Log likelihood | -3.568 | Hannan-Quinn | | 0.309 |
| F-statistic | 81.465 | DW stat | | 2.099 |
| Prob(F-statistic) | 0.000 | | | |

Table: *ARDL Model 2*

# Forecasting Euro area GDP growth

- The values of all information criteria are lower for the ARDL Model 2 compared to those for the ARDL Model 1. In addition, the $\overline{R}^2$ is slightly higher.
- This implies that the ARDL Model 2 has a better in-sample fit compared to the ARDL Model 1.
- This implies the lagged values of these two variables are more informative than their contemporaneous values.

# Forecasting Euro area GDP growth

- The reported DW test statistics for testing first order serial correlation for both ARDl Model 1 and ARDL model 2 suggest that the null hypothesis of no serial correlation cannot be rejected.
- Inclusion of the lagged variables helps capture dynamics that were present in the model errors without lagged terms.
- The Breusch-Godfrey serial correlation LM test results reported in Tables 6 and 7 show that the null hypothesis of no serial correlation up to order 2 cannot be rejected.

# Breusch-Godfrey serial correlation LM test for ARDL Model 1

| | | | | |
|---|---|---|---|---|
| F-statistic | 1.293 | Prob. F(2,60) | | 0.282 |
| Obs*R-squared | 2.891 | Prob. Chi-Square(2) | | 0.236 |
| Variable | Coefficient | Std.Error | t-Statistic | Prob. |
| C | -0.103 | 0.093 | -1.106 | 0.273 |
| Y(-1) | 0.384 | 0.325 | 1.184 | 0.241 |
| IPR | -0.011 | 0.035 | -0.325 | 0.747 |
| IPR(-1) | -0.116 | 0.106 | -1.094 | 0.279 |
| SU | 0.009 | 0.018 | 0.523 | 0.603 |
| SU(-1) | 0.007 | 0.018 | 0.389 | 0.699 |
| SR | -0.001 | 0.005 | -0.207 | 0.836 |
| SR(-1) | -0.004 | 0.006 | -0.650 | 0.518 |
| RESID(-1) | -0.481 | 0.366 | -1.315 | 0.193 |
| RESID(-2) | 0.004 | 0.164 | 0.022 | 0.982 |
| R-squared | 0.041 | Mean dep var | | 0.000 |
| Adjusted R-squared | -0.102 | S.D. dep var | | 0.253 |
| S.E. of regression | 0.266 | Akaike IC | | 0.319 |
| Sum squared resid | 4.239 | Schwarz IC | | 0.641 |
| Log likelihood | -1.180 | Hannan-Quinn | | 0.447 |
| F-statistic | 0.287 | DW stat | | 2.063 |
| Prob(F-statistic) | 0.976 | | | |

Table: *Breusch-Godfrey serial correlation LM test for ARDL Model 1*

# Breusch-Godfrey serial correlation LM test for ARDL Model 2

| | | | | |
|---|---|---|---|---|
| F-statistic | 0.662 | Prob. F(2,63) | | 0.519 |
| Obs*R-squared | 1.441 | Prob. Chi-Square(2) | | 0.486 |
| | | | | |
| Variable | Coefficient | Std.Error | t-Statistic | Prob. |
| | | | | |
| C | -0.005 | 0.042 | -0.109 | 0.914 |
| Y(-1) | 0.012 | 0.075 | 0.154 | 0.878 |
| IPR | -0.004 | 0.027 | -0.155 | 0.877 |
| SU(-1) | -0.002 | 0.013 | -0.121 | 0.904 |
| SR(-1) | 0.000 | 0.005 | 0.036 | 0.971 |
| RESID(-1) | -0.060 | 0.146 | -0.409 | 0.684 |
| RESID(-2) | 0.135 | 0.134 | 1.002 | 0.320 |
| | | | | |
| R-squared | 0.021 | Mean dep var | | 0.000 |
| AdjustedR-squared | -0.073 | S.D. dep var | | 0.256 |
| S.E. of regression | 0.266 | Akaike IC | | 0.281 |
| Sum squared resid | 4.445 | Schwarz IC | | 0.506 |
| Log likelihood | -2.840 | Hannan-Quinn | | 0.370 |
| F-statistic | 0.221 | DW stat | | 2.025 |
| Prob(F-statistic) | 0.969 | | | |

Table: *Breusch-Godfrey serial correlation LM test for ARDL Model 2*

# Forecasting with dynamic models

- We now consider forecasts of ARDL Model 2 to see if there are improvements when compared to models without dynamic, Model 2 and Model 2.3.
- We also add dummy variables to ARDL Model 2 for the Euro area crisis and the early 2000's recession created in the previous chapter.
- Table 8 presents the estimation output for the ARDL Model 2 with those dummies.
- The estimation sample is 1996Q1 to 2010Q4
- The forecast evaluation period is 2011Q1 to 2013Q2.

# ARDL Model 2 with dummies

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 0.333 | 0.055 | 6.049 | 0.000 |
| Y(-1) | 0.108 | 0.074 | 1.462 | 0.150 |
| IPR | 0.254 | 0.026 | 9.928 | 0.000 |
| SU(-1) | -0.014 | 0.013 | -1.062 | 0.293 |
| SR(-1) | 0.009 | 0.005 | 1.853 | 0.070 |
| D_EA | -0.251 | 0.089 | -2.828 | 0.007 |
| D_2000S | 0.101 | 0.096 | 1.046 | 0.300 |
| | | | | |
| R-squared | 0.874 | Mean dep var | | 0.414 |
| Adjusted R-squared | 0.860 | S.D. dep var | | 0.644 |
| S.E. of regression | 0.241 | Akaike IC | | 0.102 |
| Sum squared resid | 3.080 | Schwarz IC | | 0.346 |
| Log likelihood | 3.942 | Hannan-Quinn | | 0.198 |
| F-statistic | 61.369 | DW stat | | 2.268 |
| Prob(F-statistic) | 0.000 | | | |

Table: *ARDL Model 2 with dummies*

- Table 9 presents the forecast evaluation statistics of the three models.
- It is clear that the ARDL Model 2, having included the lagged terms, outperforms the other two models whether the forecasts are static or recursive.

# Forecast estimation statistics: 2011Q1-2013Q2

| | Model 2 | Model 2 with dummies | ARDL Model 2 with dummies |
|---|---|---|---|
| | | Static forecasts | |
| RMSFE | 0.413 | 0.326 | 0.297 |
| MAFE | 0.376 | 0.248 | 0.236 |
| | | Recursive forecasts | |
| RMSFE | 0.398 | 0.332 | 0.290 |
| MAFE | 0.363 | 0.256 | 0.236 |

Table: *Forecast evaluation statistics: 2011Q1 through 2013Q2*

# Forecasting US GDP growth

- We return to the example of forecasting US GDP growth.
- The first model we consider, the ARDL Model 1, has lags of all variables, including the dependent variable.
- The estimation sample is 1985Q1 to 2013Q4.

# ARDL Model 1 estimation

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 0.532 | 0.068 | 7.799 | 0.000 |
| Y(-1) | -0.050 | 0.086 | -0.577 | 0.565 |
| IPR | 0.308 | 0.054 | 5.668 | 0.000 |
| SU(-1) | -0.011 | 0.007 | -1.585 | 0.116 |
| SR(-1) | 0.007 | 0.008 | 0.848 | 0.399 |
| Dfincris | -0.503 | 0.256 | -1.967 | 0.052 |
| D2000s | 0.000 | 0.167 | 0.001 | 1.000 |
| | | | | |
| R-squared | 0.521 | Mean dep var | | 0.657 |
| Adjusted R-squared | 0.495 | S.D. dep var | | 0.600 |
| S.E. of regression | 0.427 | Akaike IC | | 1.192 |
| Sum squared resid | 19.833 | Schwarz IC | | 1.359 |
| Log likelihood | -62.156 | Hannan-Quinn | | 1.260 |
| F-statistic | 19.777 | DW stat | | 1.852 |
| Prob(F-statistic) | 0.000 | | | |

Table: *Estimation output: ARDL Model 2 with dummy variables*

# Forecasting US GDP growth

- Although all the lagged variables are statistically insignificant, there are still gains in forecasting from the ARDL model.
- When compared to similar static models, the adjusted $R^2$ increases substantially.
- We now estimate a simplified version of ARDL Model 1, called ARDL Model 2.

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 0.187 | 0.039 | 4.747 | 0.000 |
| Y(-1) | 0.273 | 0.065 | 4.174 | 0.000 |
| IPR | 0.262 | 0.027 | 9.827 | 0.000 |
| SU(-1) | -0.025 | 0.012 | -2.000 | 0.050 |
| SR(-1) | 0.008 | 0.005 | 1.673 | 0.099 |
| | | | | |
| R-squared | 0.834 | Mean dep var | | 0.350 |
| Adjusted R-squared | 0.823 | S.D. dep var | | 0.629 |
| S.E. of regression | 0.264 | Akaike IC | | 0.245 |
| Sum squared resid | 4.538 | Schwarz IC | | 0.405 |
| Log likelihood | -3.568 | Hannan-Quinn | | 0.309 |
| F-statistic | 81.465 | DW stat | | 2.099 |
| Prob(F-statistic) | 0.000 | | | |

Table: *ARDL Model 2*