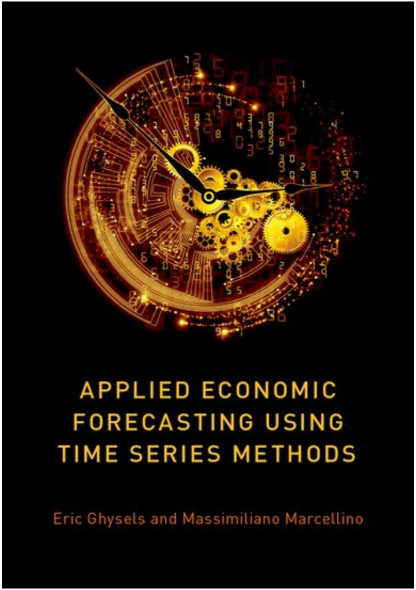

APPLIED ECONOMIC
FORECASTING USING
TIME SERIES METHODS
Eric Ghysels and Massimiliano Marcellino

Applied Economic Forecasting using Time Series Methods

Eric Ghysels and Massimiliano Marcellino

Companion Slides - Chapter 2 Model Mis-Specification

- OLS estimators are consistent, efficient, and asymptotically normally distributed if the following assumptions (1.2.1) hold.

Assumption (Linear Regression Assumptions)

In the regression model in equation (1.2.1):

- *[LR1]* $E(\varepsilon) = 0$,
- *[LR2]* $E(\varepsilon\varepsilon') = \sigma^2 I_T$,
- *[LR3]* X *is distributed independently of* ε,
- *[LR4]* $X'X$ *is non singular,*
- *[LR5]* X *is weakly stationary.*

- What happens if each of these hypotheses is violated, how we can assess if the hypotheses hold or not, and what we can do if they do not?

Heteroskedastic and correlated errors

- The linear regression model with k explanatory variables can be written in compact notation as

$$\underset{T\times 1}{y} = \underset{T\times k}{X}\underset{k\times 1}{\beta} + \underset{T\times 1}{\varepsilon}, \tag{1}$$

Assumption (Linear Regression Assumptions)

In the regression model in equation (2.1.1):

- *[LR2a] The errors ε are homoskedastic, i.e., Var$(\varepsilon_i) = \sigma^2$, $i = 1, \ldots, T$,*
- *[LR2b] The errors ε are uncorrelated, i.e., $corr(\varepsilon_t, \varepsilon_{t-j}) = 0$, $j = 1, \ldots$*

- It is probable that there is heteroskedasticity, i.e., Var$(\varepsilon_i) = \sigma_i^2$, $i = 1$, $\ldots$,T. The error variance changes with the different observations. The errors can also be correlated over time, i.e., $corr(\varepsilon_t, \varepsilon_{t-j}) \neq 0$. More specifically, let us consider the case of first order serial correlation, $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$.

Heteroskedastic and correlated errors

- A simple way to formalize heteroskedasticity and serial correlation within the linear model is by changing the representation of the error's variance-covariance matrix from $\sigma^2 I$ to simply Ω. For example, if errors are uncorrelated but their variance changes across observations, Ω will be:

$$\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \sigma_T^2 \end{bmatrix} \tag{2}$$

Heteroskedastic and correlated errors

- Instead, if the errors display first order serial correlation and $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$, $u_t \overset{iid}{\sim} (0, \sigma_u^2)$, then we have:

$$\Omega = \sigma_u^2/(1-\rho^2) \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & & \vdots \\ \rho^2 & \rho & & & \rho^2 \\ \vdots & & & \ddots & \rho \\ \rho^{T-1} & \dots & \rho^2 & \rho & 1 \end{bmatrix}. \quad (3)$$

- When Assumptions 2 - LR2a-LR2b are substituted with

$$\text{Var}(\varepsilon) = \Omega, \quad (4)$$

the resulting model is known as the generalized linear model.

Heteroskedastic and correlated errors

Consequences of violating the assumptions of homoskedastic and uncorrelated errors:

- First, the OLS estimator is no longer efficient.
- Second, $\hat{\beta}_{OLS}$ is still consistent but the formula for its variance covariance matrix requires modifications. Moreover, consistency is lost when the model is dynamic and the errors are serially correlated.
- Third, the variance estimator $\hat{\sigma}^2_{OLS} = \hat{\varepsilon}'\hat{\varepsilon}/(T-k)$ will be biased and inconsistent.
- Fourth, the standard versions of the confidence intervals and of the $t-$ and $F-$tests, which rely on the variance of $\hat{\beta}_{OLS}$, are no longer valid.

Heteroskedastic and correlated errors

Remedies to the problems:

- Improve the model specification when there is evidence that there are problems like omitted variables or model instability. e.g. Using a dynamic rather than a static model.
- Make appropriate transformations of the variables. e.g. Working with variables in logarithms, working with variables expressed in growth rates instead of levels.
- Change the method to compute the variance estimator for $\hat{\beta}_{OLS}$. e.g.The Heteroskedasticity and Autocorrelation Consistent (HAC) variance estimator exploits.
- Change the estimation method to GLS.

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator

- Since the error covariance matrix Ω is positive definite, there exists an invertible matrix H such that

$$H\Omega H' = I.$$

- Hence

$$\Omega = H^{-1}(H')^{-1} = (H'H)^{-1}$$

- Let us consider again the linear model (1): $y = X\beta + \varepsilon$ where now $\mathrm{Var}(\varepsilon) = \Omega$. If we multiply both sides by H we obtain

$$Hy = HX\beta + H\varepsilon$$

or

$$\tilde{Y} = \tilde{X}\beta + \tilde{\varepsilon} \tag{5}$$

where

$$E(\tilde{\varepsilon}\tilde{\varepsilon}') = H\Omega H' = I.$$

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator

- It is now possible to use OLS in the transformed model (5), hence obtaining the GLS estimator:

$$\hat{\beta}_{GLS} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y. \tag{6}$$

  Let us describe the properties of the GLS estimator:
  - $\hat{\beta}_{GLS}$ is unbiased as:

$$E(\hat{\beta}_{GLS}) = E((\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y}) = \beta + E((\tilde{X}'\tilde{X})^{-1}\tilde{X}'E(\tilde{\varepsilon})) = \beta.$$

  - The variance of $\hat{\beta}_{GLS}$ is:

$$Var(\hat{\beta}_{GLS}) = E(\hat{\beta}_{GLS} - \beta)(\hat{\beta}_{GLS} - \beta)' = (X'\Omega^{-1}X)^{-1}.$$

  - $\hat{\beta}_{GLS}$ is efficient, Gauss Markov Theorem.
  - $\hat{\beta}_{GLS}$ is consistent.
  - The asymptotic distribution of $\hat{\beta}_{GLS}$ is normal,

$$\hat{\beta}_{GLS} \overset{a}{\sim} N(\beta, (X'\Omega^{-1}X)^{-1}).$$

  - If $\Omega = \sigma^2 I$, then $\hat{\beta}_{GLS} = \hat{\beta}_{OLS}$.

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator

- In particular, Ω is not known and has to be estimated.The GLS estimator where the Ω matrix is substituted by its estimator $\hat{\Omega}$ takes the name of Feasible GLS (FGLS) estimator:

$$\hat{\beta}_{FGLS} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}Y.$$

- Estimation of the Ω matrix is not simple, since in general it consists of $T(T+1)/2$ distinct unknown elements while the number of observation is only T.
- For this reason, one needs to impose some a priori restrictions on Ω.

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator

- Let us take the case of heteroskedasticity. Assuming that the variance is equal to σ_1^2 in a first subsample of length T_1 while it becomes σ_2^2 in a second subsample of length T_2, we can estimate both σ_1^2 and σ_2^2 using the standard OLS formula applied separately in each subsample. The resulting $\hat{\Omega}$ will be:

$$\hat{\Omega} = \begin{bmatrix} \hat{\sigma}_1^2 & \dots & 0 & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ 0 & \dots & \hat{\sigma}_1^2 & 0 & \dots & 0 \\ 0 & \dots & 0 & \hat{\sigma}_2^2 & \dots & 0 \\ \vdots & & \vdots & & \ddots & \\ 0 & \dots & 0 & 0 & \dots & \hat{\sigma}_2^2 \end{bmatrix} \tag{7}$$

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator

- In the case of serially correlated errors, one can estimate by OLS the original regression model and use the resulting residuals $\hat{\varepsilon}_t$ to estimate the model $\hat{\varepsilon}_t = \rho\hat{\varepsilon}_{t-1} + u_t$. Then, the OLS estimators of ρ and Var(u_t), $\hat{\rho}$ and $\hat{\sigma}_u^2$, can be used to build $\hat{\Omega}$ as:

$$\hat{\Omega} = \frac{\hat{\sigma}_u^2}{1-\hat{\rho}^2}\begin{bmatrix} 1 & \hat{\rho} & \hat{\rho}^2 & \dots & & \hat{\rho}^{T-1} \\ \hat{\rho} & 1 & \hat{\rho} & \dots & & \vdots \\ \hat{\rho}^2 & \hat{\rho} & & & & \hat{\rho}^2 \\ \vdots & & & & & \hat{\rho} \\ \hat{\rho}^{T-1} & & \dots & \hat{\rho}^2 & \hat{\rho} & 1 \end{bmatrix}. \tag{8}$$

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator

- Since we used a regression of y on X to obtain $\hat{\Omega}$, $\hat{\Omega}$ and ε are in general correlated.
- The FGLS estimator is biased in small samples, since

$$E\left[(X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}\varepsilon\right] \neq 0 \Rightarrow E(\hat{\beta}_{FGLS}) \neq \beta. \tag{9}$$

- $\hat{\beta}_{FGLS}$ is no longer a linear estimator, and it is not necessarily the minimum variance estimator.
- However, $\hat{\beta}_{FGLS}$ remains a consistent estimator, $\hat{\beta}_{FGLS} \rightarrow \beta$, and asymptotically it has the same properties as $\hat{\beta}_{GLS}$.

The Generalized Least Squares (GLS) estimator and the feasible GLS estimator

- In terms of forecasting, it can be shown that the optimal (in the MSFE sense) h-steps ahead forecast is

$$\hat{y}_{T+h} = x_{T+h}\hat{\beta}_{GLS} + W'\Omega^{-1}\hat{\varepsilon}, \tag{10}$$

where $W = E(\varepsilon_{T+h}, \varepsilon)$, which is a $1 \times T$ vector containing as elements $E(\varepsilon_{T+h}, \varepsilon_t) \; \forall \; t = 1, \ldots, T$, see, e.g., Granger and Newbold (1986,p.191).

HAC estimators

- In the section, we deal with the estimation of variances in settings where we do not specify a (parametric) model for the process involved.
- Situation one: Heteroskedasticity.
  - Suppose the regression errors ε_i are independent, but have distinct variances σ_i^2, $i = 1, \ldots, T$. Then $\Omega = diag(\sigma_1^2, \ldots, \sigma_T^2)$, and $\hat{\sigma}_i^2$ can be estimated with $\hat{\varepsilon}_i^2$, yielding $\hat{\Omega} = diag(\hat{\varepsilon}_1^2, \ldots, \hat{\varepsilon}_T^2)$.
  - It can then be shown that using an expression for Var$(\hat{\beta}_{OLS})$ similar to $(X'X)^{-1}X'\hat{\Omega}X(X'X)^{-1}$ instead of $\hat{\sigma}^2(X'X)^{-1}$ yields White's estimator, often referred to as a heteroskedasticity consistent estimator.

HAC estimators

- Situation two: Correlated errors.
  - Let us define $\hat{\Gamma}_{j,T} = (1/(T-j)) \sum_{t=j+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}'_{t-j}$. Then, we can consider the following estimator for $var(\varepsilon_t)$, which we will denote by $\hat{V}_T$:

$$\hat{V}_T = \hat{\Gamma}_{0,T} + \left[\sum_{j=1}^{T-1} w_{j,T} (\hat{\Gamma}_{j,T} + \hat{\Gamma}'_{j,T}) \right] \quad (11)$$

    where $w_{j,T}$ is a weighting scheme, also known as kernel.
  - The procedure is usually referred to as Heteroskedasticity and Autocorrelation Consistent (HAC) variance estimation. See Newey and West(1987) and Andrews(1991).

- The details of the kernel functions and bandwidth appear in the table below.

Bartlett, Parzen and Quadratic Spectral (QS) kernels

Kernel	$w_{j,T}$	Optimal Bandwith
Bartlett	$k(j,b_T) = \max(1-a_j,0) \quad a_j = j/(1+b_T)$	$O(T^{1/3})$
Parzen	$k(j,b_T) = \begin{cases} 1-6a_j^2+6a_j^3 & \text{for } 0 \leq a_j \leq 1/2 \\ 2(1-a_j)^3 & \text{for } 1/2 < a_j \leq 1 \\ 0 & \text{otherwise} \end{cases}$	$O(T^{1/5})$
QS	$k(j,b_T) = \frac{25}{12\pi^2 d_j^2}\left(\frac{\sin(m_j)}{(m_j)} - \cos(m_j)\right)$ $d_j = j/b_T \quad m_j = 6\pi d_j/5$	$O(T^{1/5})$

Some tests for homoskedasticity and no correlation

- A first heuristic method to check for homoskedasticity is to plot $\hat{\varepsilon}_i^2, i = 1, \ldots, T$, where as usual we indicate with $\hat{\varepsilon}$ the residuals of the OLS regression, $\hat{\varepsilon} = y - X'\hat{\beta}_{OLS}$. If the variability of $\hat{\varepsilon}_i^2$ changes substantially over time (or across observations) this might be a sign of heteroskedasticity.
- In this section, we will consider some more rigorous testing procedure:
  Goldfeld-Quandt test
  Breusch-Pagan-Godfrey
  White tests for homoskedasticity

- Goldfeld-Quandt test assumes the null hypothesis $H_0 : \sigma_i^2 = \sigma^2$, while the alternative is $H_1 : \sigma_i^2 = cz_i^2$, $c > 0$, (the variance increases with the explanatory variable z.)
  - Step 1: the sample observations (y_i, x_i), $i = 1, \ldots, T$, are ranked according to the values of z_i, so that the lowest values of x_i are in the first part of the sample.
  - Step 2: the d central observations of the re-ordered sample are excluded (e.g. 20% of the sample, or $d = 0.2T$), and the remaining observations are divided in two subsamples of size $(T - d)/2$.
  - Step 3: for each subsample one computes the Residual Sum of Squares $RSS = \sum \hat{\varepsilon}_i^2$, and constructs the following ratio:

$$GQ = \frac{RSS_2}{RSS_1}, \tag{12}$$

  where RSS_j refers to subsample $j = 1$ and 2.

Under the null hypothesis the test statistic GQ is distributed as $F(p, p)$, where $p = [(T - d)/2]$- k degrees of freedom.

- Breusch-Pagan-Godfrey test assumes the null hypothesis $H_0 : \sigma_i^2 = \sigma^2$, under the alternative there is an unknown relationship between errors' variance and Z, i.e. $H_1 : \sigma_i^2 = \gamma + \delta Z_i$.
  - Step 1: using the OLS residuals $\hat{\varepsilon}$ one regresses the squared residuals on Z:
  $$\hat{\varepsilon}_i^2 = \gamma + \delta Z_i + v_i. \tag{13}$$
  - Step 2: one computes the Breusch-Pagan-Godfrey (BPG) test as:
  $$BPG = TR^2, \tag{14}$$
  where R^2 is simply the coefficient of determination of regression (13), and T is the sample size.

  Under the null hypothesis, and the basic assumptions of the linear regression model, asymptotically $BPG \sim \chi_q^2$, where q is the dimension of Z.

White's homoskedasticity test

- White's homoskedasticity test assumes the null hypothesis $H_0 : \sigma_i^2 = \sigma^2$, while the alternative is $H_1 : \sigma_i^2 = f(X_i, Z_i)$.
  - Step 1: we run the regression (assuming for simplicity that X and Z are scalar):
  $$\hat{\varepsilon}_i^2 = \gamma_0 + \gamma_1 X_i + \gamma_2 Z_i + \gamma_3 X_i^2 + \gamma_4 Z_i^2 + \gamma_5 X_i Z_i + v_i \tag{15}$$
  In the event of multiple regressors in Z, we take all linear, quadratic and cross-product terms.
  - Step 2: similar to the BPG test, we compute the White's statistic (W) as: W= TR^2 for regression (15).

  Under the null, the asymptotic distribution is again χ_q^2, where q is the number of regressors (other than the constant) in the above regression.

Some tests for homoskedasticity and no correlation

- Durbin-Watson (DW) test for no serial correlation. The null hypothesis is $H_0 : \varepsilon_t \ uncorrelated$, while the alternative is first order correlation, i.e. $H_1 : \varepsilon_t = \rho\, \varepsilon_{t-1} + u_t, \ \ \rho \neq 0$. The test statistic is based on the OLS residuals:

$$DW = \frac{\sum_{t=2}^{T} (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{T} \hat{\varepsilon}_t^2} \approx 2 - 2\hat{\rho}. \tag{16}$$

  Hence $0 \leq DW \leq 4$, with $DW \approx 2$ under the null.

- Under the null H_0 the test statistic DW has a non-standard distribution, and it is not possible to use it with dynamic specifications. Moreover, user might be interested in detecting also serial correlation of a order higher than one.

Some tests for homoskedasticity and no correlation

- Lagrange Multipliers (LM) test, also called the Breusch-Godfrey test, is based on the null hypothesis $H_0 : \varepsilon_t \; uncorrelated$, while the alternative is $H_1 : \varepsilon_t \; correlated \; up \; to \; order \; m$, where m is user-specified. Defining

$$\hat{r}_j = \frac{\sum_{t=j+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}}{\sum_{t=1}^{T} \hat{\varepsilon}_t^2}, \tag{17}$$

for $j = 1, \ldots, m$, the test statistic is:

$$LM = T \left(\sum_{j=1}^{m} \hat{r}_j^2 \right). \tag{18}$$

Under H_0, LM is asymptotically distributed as χ^2_m, assuming again that all the other model assumptions hold.

Parameter instability

In this section, we will discuss the following topics:

- Sources of parameter instability
- Parameter instability in linear regression model
- Implications for forecasting
- Detect parameter instability
- Use of dummy variables as a remedy

The effects of parameter changes

- Let us reconsider the basic model and allow for time-varying parameters:
$$y_t = X_{1t}\beta_{1t} + X_{2t}\beta_{2t} + \ldots X_{kt}\beta_{kt} + \varepsilon_t, \tag{19}$$
- We have so far imposed the following:
[LR6] The parameters are stable across time, $\beta_{it} = \beta_i \ \forall \ i$ and t.
- Now suppose that at time T_1 a potentially destabilizing event happens. The model can be rewritten as:

$$y_t = X_t\beta_1 + \varepsilon_{1t} \quad t = 1, \ldots, T_1 \text{ or } t \in \mathcal{T}_1, \tag{20}$$

$$y_t = X_t\beta_2 + \varepsilon_{2t} \quad t = T_1 + 1, \ldots, T \text{ or } t \in \mathcal{T}_2, \tag{21}$$

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \overset{iid}{\sim} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 I_{T_1} & 0 \\ 0 & \sigma_2^2 I_{T_2} \end{bmatrix} \right). \tag{22}$$

The effects of parameter changes

- If we assume that the parameters are constant by mistake, then the OLS estimators $\hat{\beta}$ and $\hat{\sigma}_u^2$ are biased and inconsistent.
- Moreover, forecasts of future values of y is clearly suboptimal. Suppose that $\mathcal{T}_1$ is the estimation sample and $\mathcal{T}_2$ the forecast sample.
  Assuming that future values of X_t are known, the optimal forecast made in $\mathcal{T}_1$ for y_t over $t \in \mathcal{T}_2$ is

$$\hat{y}_t = X_t \beta_1.$$

  The actual values are instead:

$$y_t = X_t \beta_2 + \varepsilon_{2t},$$

  so that the forecast errors are:

$$e_t = X_t(\beta_2 - \beta_1) + \varepsilon_{2t}.$$

  Therefore, the larger the change in β the larger the forecast error (and the MSFE).

Simple tests for parameter changes

- Let us consider the following sets of hypotheses:

$$H_0' : \beta_1 = \beta_2 \ , \quad H_0'' : \sigma_1^2 = \sigma_2^2, \tag{23}$$
$$H_1' : \beta_1 \neq \beta_2 \ , \quad H_1'' : \sigma_1^2 \neq \sigma_2^2. \tag{24}$$

  and where $H_0 = H_0' \cup H_0''$.
- Denote $T2$ as the number of observations in the second subsample period, k as the number of regressors. It is convenient to consider two separate cases:
- Case 1: $T_2 > k$
- Case 2: $T_2 < k$

Simple tests for parameter changes

- Case 1: $T_2 > k$
  It is possible to estimate both models on the whole set of observations, and the two models in the respective subsamples $\mathcal{T}_1$ and $\mathcal{T}_2$. Three sets of residuals sum of squares (RSS), denoted by RSS_T, RSS_{T_1}, and RSS_{T_2}.
  The Chow test for the null hypothesis $H_0' : \beta_1 = \beta_2 | H_0''$ is

$$CH_1 : (\frac{RSS_T - RSS_{T_1} - RSS_{T_2}}{RSS_{T_1} + RSS_{T_2}}) \cdot \frac{T - 2k}{k} \underset{H_0}{\sim} F(k, T - 2k), \tag{25}$$

where $T = T_1 + T_2$

Simple tests for parameter changes

- The null hypothesis for CH_1 is built on the simplifying assumption that the variance remains constant across the two subsamples H_0' : $\beta_1 = \beta_2 | H_0''$.
- Therefore, we would like to test:

$$H_0'' : \sigma_1^2 = \sigma_2^2 \text{ against } H_1'' : \sigma_1^2 < \sigma_2^2. \tag{26}$$

The test statistic is now

$$CH_2 = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} = \frac{RSS_{T_2}}{RSS_{T_1}} \frac{T_1 - k}{T_2 - k} \underset{H_0''}{\sim} F(T_2 - k, T_1 - k). \tag{27}$$

It can be shown that CH_1 and CH_2 are independent so one can first use CH_2 to verify the variance stability and then to apply CH_1.
- Compare CH_2 test with homoskedasticity test.

- Case 2: $T_2 < k$

  In this case there are not enough degrees of freedom to estimate model (21). The null $H_0' : \beta_1 = \beta_2 | H_0''$ must then be tested through a different test, which is:

$$CH_3 : \left(y_2 - X_2\hat{\beta}_1\right)' \frac{\left[I + X_2(X_1'X_1)^{-1}X_2'\right]^{-1}}{\hat{\varepsilon}_1'\hat{\varepsilon}_1} \left(\underset{T_{2x1}}{y_2} - X_2\hat{\beta}_1\right) \frac{T_1 - k}{T_2}$$

$$\underset{H_0}{\sim} F(T_2, T_1 - k).$$

y_2 - $X_2\hat{\beta}_1$ is the forecast error.

$$\left(\left[I + X_2\left(X_1'X_1\right)^{-1}X_2'\right]^{-1} / \hat{\varepsilon}_1'\hat{\varepsilon}_1\right)(T_1 - k)$$

is an estimator of the forecast errors' covariance matrix.

Simple tests for parameter changes

- We can rewrite the CH_3 test:

$$CH_3 = \left(\frac{RSS_T - RSS_{T_1}}{RSS_{T_1}} \right) \frac{T_1 - k}{T_2} \underset{H_0}{\sim} F(T_2, T_1 - k). \qquad (28)$$

- To test variance stability, i.e. $H_0'' : \sigma_1^2 = \sigma_2^2$ against $H_1'' : \sigma_1^2 < \sigma_2^2$. The test statistics is now

$$CH_4 : \frac{\left(y_2 - X_2\hat{\beta}_1\right)' \left(y_2 - X_2\hat{\beta}_1\right)}{\hat{\sigma}_1^2} \underset{H_0''}{\overset{a}{\sim}} \chi^2(T_2). \qquad (29)$$

- The numerator and denominator of CH_4 are not independent. Therefore, the distribution of CH_4 is not F-shaped.

Simple tests for parameter changes

- Chow tests we have discussed require the specification of the date of the parameter change.
- If this is not known with certainty, we can compute the tests for a set of possible candidate break dates, and then take the maximum (or supremum) of the resulting set of statistics.
- Unfortunately, the max-Chow test does not have a standard distribution, so that the proper critical values must be tabulated using simulation methods.

- An alternative to sophisticated versions of the Chow tests in the presence of unknown break points is recursive estimation.
- The general form of the recursive OLS estimator underlying the recursive forecasts we have discussed in Chapter 1 is:

$$\hat{\beta}_t = \Big(\underset{k\times t}{\bar{X}_t'}\bar{X}_t\Big)^{-1} \bar{X}_t' \underset{t\times 1}{\bar{y}_t} \qquad t = T_0, T_0+1, \ldots, T \tag{30}$$

The corresponding estimator for the variance of $\hat{\beta}_t$ is

$$var\left(\hat{\beta}_t\right) = \hat{\sigma}_t^2 \Big(\underset{k\times t}{\bar{X}_t'}\bar{X}_t\Big)^{-1} \tag{31}$$

with

$$\hat{\sigma}_t^2 = (\bar{y}_t - \bar{X}_t\hat{\beta}_t)'(\bar{y}_t - \bar{X}_t\hat{\beta}_t)/(T-k). \tag{32}$$

Recursive methods

- Plotting the recursive estimators of the parameters $\hat{\beta}_t$, $t = T_0, T_0 + 1, \ldots, T$ with their confidence bands can prove useful in detecting potential parameter instability.
- If β is constant over time, then $\hat{\beta}_t$ should quickly settle down to a common value.
- If for example the true model was (20) and (21), then we would expect to see $\hat{\beta}_t$ starting very close to β_1 and, after having trespassed $t = T_1 + 1$, to get closer and closer to β_2.

Recursive methods

Formal tests for parameter instability can be computed from the one-step ahead recursive residuals.

- One-step ahead forecast errors, given by:

$$\tilde{\varepsilon}_t = \left(y_t - X_t\hat{\beta}_{t-1}\right) = \varepsilon_t + X_t(\beta - \hat{\beta}_{t-1}) \quad (33)$$

  with

$$\hat{\beta}_{t-1} = \Bigg(\underset{k\times(t-1)}{\bar{X}_{t-1}}\ \bar{X}_{t-1}\Bigg)^{-1} \bar{X}'_{t-1}\ \underset{(t-1)\times 1}{\bar{y}_{t-1}} \quad (34)$$

  and

$$\tilde{\sigma}_t^2 = Var(\tilde{\varepsilon}_t) = (1 + X_t(\bar{X}'_{t-1}\bar{X}_{t-1})^{-1}X'_t)\sigma^2. \quad (35)$$

Recursive methods

- The standardized recursive residuals can be defined as:

$$\tilde{\omega}_t = \tilde{\varepsilon}_t / \tilde{\sigma}_t, \qquad t = k+1, \ldots, T \tag{36}$$

- Brown, Durbin and Evans (1975) propose a CUSUM statistic:

$$CUSUM_t = \sum_{j=k+1}^{t} \frac{\tilde{\omega}_j}{\tilde{\sigma}_{\omega,T}} \tag{37}$$

where $\tilde{\sigma}^2_{\omega,T} = (T-k)^{-1} \sum_{t=k+1}^{T} (\tilde{\omega}_t - \bar{\omega})^2$.

- Under the null hypothesis that $\beta_{k+1} = \ldots = \beta_T$, $CUSUM_t$ has mean zero and variance that is proportional to $t - k - 1$. The CUSUMSQ statistic is defined as:

$$CUSUMSQ_t = \frac{\sum_{j=k+1}^{t} \tilde{\omega}_j^2}{\sum_{j=k+1}^{T} \tilde{\omega}_j^2} \tag{38}$$

- Under the null of parameter stability, the $CUSUMSQ_t$ statistic has an asymptotic $\chi^2(t)$ distribution.

Dummy variables

- Suppose the basic consumption model is

$$c_t = \beta_1 + \beta_2 inc_t + \varepsilon_t,$$

where c indicates consumption and inc income. Let us insert a dummy variable D_t taking value one during credit crunch times and zero for the remaining part of the sample.

- We focus on the model that accounts for structural changes in both the intercept and the slope of consumption function:

$$c_t = \beta_1 + \alpha D_t + \beta_2 inc_t + \gamma D_t inc_t + \varepsilon_t$$

- Assume, for example, that at time t_0 there is a permanent shift in the marginal propensity to consume. The model could then be specified as

$$c_t = \beta_1 + \beta_2 inc_t + \beta_3 (inc_t - inc_{t_0}) D_t + \varepsilon_t \tag{39}$$

$$D_t = \begin{cases} 1 \text{ if } & t \geq t_0 \\ 0 \text{ if } & t < t_0 \end{cases} \tag{40}$$

- Using dummy variables, we can also rewrite model (20)-(21) as a single equation. Assuming that the error variance is constant, it is:

$$y_t = \beta_1 X_t + \gamma_2 D_t X_t + \varepsilon_t, \tag{41}$$

with $\beta_2 = \beta_1 + \gamma_2$ and

$$D_t = \begin{cases} 1 \text{ if } & t > T_1 \\ 0 \text{ if } & t \leq T_1 \end{cases} \tag{42}$$

It can be shown that an $F-$test for the null hypothesis $\gamma_2 = 0$ in (41) is exactly equivalent to the CH_1 test in (25). Under the null, the $F-$statistic has an $F(k, T-2k)$ distribution.

- When the variance is also changing across sub-periods

$$\text{Var}(\varepsilon_t) = \begin{cases} \sigma_1^2, & \text{if } t \leq T_1 \\ \sigma_2^2, & \text{if } t > T_1 \end{cases} \tag{43}$$

the model (41) can be reformulated as

$$\begin{aligned} y_t/((1-D_t)\sigma_1 + D_t\sigma_2) &= \beta_1 X_t/((1-D_t)\sigma_1 + D_t\sigma_2) + \gamma_2 D_t X_t/ \\ &\quad ((1-D_t)\sigma_1 + D_t\sigma_2) + u_t \end{aligned} \tag{44}$$

where $u_t = \varepsilon_t/((1-D_t)\sigma_1 + D_t\sigma_2)$, $\text{Var}(u_t) = 1$.

- When the structural change is more gradual, but its shape is still known, dummies can take more general forms, such as

$$D_t = \begin{cases} 1/(1+\exp(\mu t)) & \text{if } t \geq t_0 \\ 0 & \text{if } t < t_0 \end{cases} \tag{45}$$

In more complex models, this gradual structural change may be driven by a specific regressor, as for example in

$$D_t = \begin{cases} 1/(1+\exp(\mu Z_t)) & \text{if } Z_t \geq Z_0 \\ 0 & \text{if } Z_t < Z_0 \end{cases} \tag{46}$$

In these cases, the model becomes non-linear and a non-linear least squares (NLS) estimation method is required.

Multiple breaks

- Bai and Perron (1998) propose a procedure to determine the number of breaks and their location. We consider a standard multiple linear regression model as in equation estimated over a sample of T observations:

$$y_t = X_{1t}\beta_1 + X_{2t}\beta_2 + ...X_{kt}\beta_k + \varepsilon_t.$$

  Suppose now that among the regressors, $X_{1t} \ldots, X_{kt}$ we identify a k_1-dimensional set X_t^c which is not subject to potential structural change, whereas the remainder k_2-dimensional set called Z_t might be affected by parameter instability, where $k = k_1 + k_2$.
- We have a sample of size T periods and consider m potential breaks, producing $m + 1$ sample segments with stable parameters. For the observations $T_j, T_j + 1 \ldots, T_{j+1} - 1$ in regime j we have the regression model

$$y_t = X_t^c\beta + Z_t\alpha_j + \varepsilon_t, \quad (47)$$

  for regimes $j = 0, \ldots, m$.

Multiple breaks

- Bai and Perron (1998) propose a test for equality of the α_j across regimes. The null hypothesis is $H_0 : \alpha_0 = \ldots = \alpha_m$ and while the alternative hypothesis is that there are m breaks. The associated $F-$statistic:

$$F_m(\hat{\alpha}) = \frac{1}{T}\left(\frac{T-(m+1)k_2-k_1}{mk_2}\right)(R\hat{\alpha})'(R\hat{V}(\hat{\alpha})R')^{-1}(R\hat{\alpha}) \quad (48)$$

where $\hat{\alpha} = (\hat{\alpha}_0 \ldots, \hat{\alpha}_m)$ and $(R\hat{\alpha})' = (\alpha_0' - \alpha_1', \ldots, \alpha_{m-1}' - \alpha_m')$ and $\hat{V}(\hat{\alpha})$ is the estimated variance matrix of $\hat{\alpha}$.

Multiple breaks

- Where the number of breaks m is not known, we may test the null of no structural change against an unknown number of breaks up to some upper-bound m^*.
- This type of testing is sometimes termed double maximum since it involves optimization both for a given m and across various values of the test statistic for $m \leq m^*$.
- The resulting statistics is sometimes referred to as $supF_m$. Both F_m and $supF_m$ have non-standard asymptotic distributions. Bai and Perron (1998) provide critical value and response surface computations for various trimming parameters, number of regressors and number of breaks.

Measurement error and real-time data

- For real-time data, sometimes Assumption 1.2.1 - LR3 is violated, namely, the explanatory variables X are stochastic and not distributed independently of ε.
- For example, GDP for quarter T can be released in preliminary form (so called flash estimates) as soon as 30 days into the next quarter $T + 1$ and updated frequently thereafter up to several years later (see e.g, Ghysels, Horan, and Moench 2017).
- Econometrician is bounded to use real-time data.

Measurement error and real-time data

Let us assume that $E(\varepsilon|X) \neq 0$, so that ε and X are correlated. Then,

- The OLS estimator for β is biased:

$$E(\hat{\beta}) = \beta + E_{f(\varepsilon,X)}[(X'X)X'\varepsilon] = \beta + E_{f(x)}[(X'X)X']E_{f(\varepsilon|X)}(\varepsilon|X) \neq \beta,$$

- $\hat{\beta}$ and $\hat{\sigma}^2$ are inconsistent:

$$\begin{aligned}
\hat{\beta} &= \beta + (X'X)^{-1}X'\varepsilon = \beta + \left(\frac{X'X}{T}\right)^{-1}\frac{X'\varepsilon}{T} \\
&\overset{T\to\infty}{\to} \beta + (\Sigma_{XX}^{-1})\text{plim}\frac{X'\varepsilon}{T} \neq \beta, \\
\hat{\sigma}^2 &= \frac{\hat{\varepsilon}'\hat{\varepsilon}}{T-k} = \frac{\varepsilon'[I - X(X'X)X']\varepsilon}{T-k} \\
&= \left[\frac{\varepsilon'\varepsilon}{T} + \frac{\varepsilon'X}{T}\left(\frac{X'X}{T}\right)^{-1}\frac{X'\varepsilon}{T}\right]\frac{T}{T-k} \\
&\overset{T\to\infty}{\to} \sigma^2 + \text{plim}\frac{\varepsilon'X}{T}(\Sigma_{XX}^{-1})\text{plim}\frac{X'\varepsilon}{T} \neq \sigma^2.
\end{aligned}$$

- Let us illustrate the measurement error using the following example. The model is:

$$y = X\beta + \varepsilon \text{ with } Cov(X, \varepsilon) = 0 \tag{49}$$

but the observed regressors are

$$X^* = X + v,$$

where v represents the measurement error and has the following properties:

$$E(v) = 0\ ,\ \ Var(v) = \sigma_v^2 I\ ,\ \ Cov(X, v) = 0\ ,\ \ Cov(\varepsilon, v) = 0. \tag{50}$$

- If we rewrite the regression model in terms of observable variables we have:

$$Y = X^*\beta + \varepsilon - v\beta = X^*\beta + u,$$

so that $Cov(X^*, u) \neq 0$ and the OLS estimator for β, $\hat{\beta} = (X^{*\prime}X^*)^{-1}X^{*\prime}y$, is not consistent.

Measurement error and real-time data

- Let us now assess what happens when the dependent variable y is measured with error (v), while the regressors are independent of (or at least uncorrelated with) the error term. Let us suppose that

$$y^* = y + v, \tag{51}$$

with

$$E(v) = 0, \;\; Var(v) = \sigma_v^2, \;\; Cov(X, v) = 0, \;\; Cov(\varepsilon, v) = 0.$$

- The model cast in terms of observables is now

$$y = y^* - v = X\beta + \varepsilon \;\text{ or }\; y^* = X\beta + \varepsilon + v.$$

- Since $Cov(X, \varepsilon + v) = 0$, the OLS estimator, $\hat{\beta} = (X'X)^{-1} X'y^*$, remains consistent. However, its variance increases to $(\sigma_u^2 + \sigma_v^2)E(X'X)^{-1}$, resulting in a loss of efficiency.

Instrumental variables

- One way to solve the endogeneity problem is to find some exogenous variables. Let us assume there exist q variables Z with the following properties:

$$\begin{aligned} \text{plim } Z'\varepsilon/T &= 0, \\ \text{plim } Z'X/T &= \Sigma_{ZX}, \\ \text{plim } Z'Z/T &= \Sigma_{ZZ}. \end{aligned} \tag{52}$$

- Variables with these properties are called Instrumental Variables (IV). The number of instrumental variables, q, must be at least equal to the number of regressors correlated with the error term.

- The IV estimator has the following formulation:

$$\hat{\beta}_{IV} = \left(Z'X\right)^{-1} Z'y.$$

Let us multiply by Z' both sides of the model

$$y = X\beta + \varepsilon,$$

obtaining:

$$Z'y = Z'X\beta + Z'\varepsilon.$$

If we assume that (52) also holds in finite samples, so that $Z'\varepsilon = 0$, we have

$$Z'y = Z'X\hat{\beta}_{IV} \Rightarrow \hat{\beta}_{IV} = \left(Z'X\right)^{-1} Z'y.$$

Instrumental variables is an example of method of moments estimator.

Instrumental variables

- The properties of $\hat{\beta}_{IV}$ are the following:
  - $\hat{\beta}_{IV}$ is consistent:

$$\begin{aligned}\hat{\beta}_{IV} &= (Z'X)^{-1} Z'y = \beta + (Z'X)^{-1} Z'\varepsilon \\ &= \beta + \left(\frac{Z'X}{T}\right)^{-1} \left(\frac{Z'\varepsilon}{T}\right) \overset{T\to\infty}{\to} \beta + \Sigma_{zx}^{-1} \cdot 0 = \beta.\end{aligned}$$

  - Its asymptotic variance is:

$$\begin{aligned}Var\left(\hat{\beta}_{IV}\right) &= E\left[\left(\hat{\beta}_{IV} - \beta\right)\left(\hat{\beta}_{IV} - \beta\right)'\right] \\ &= E\left[(Z'X)^{-1} Z'\varepsilon\varepsilon' Z\,(Z'X)^{-1}\right] \overset{T\to\infty}{\to} \frac{\sigma^2}{T}\Sigma_{ZX}^{-1}\Sigma_{ZZ}\Sigma_{ZX}^{-1}.\end{aligned}$$

  - The asymptotic distribution of $\hat{\beta}_{IV}$ is given by:

$$\sqrt{T}\left(\beta - \hat{\beta}_{IV}\right) \overset{a}{\sim} N(0, \sigma^2\Sigma_{ZX}^{-1}\Sigma_{ZZ}\Sigma_{ZX}^{-1}).$$

Instrumental variables

- Note $Var(\hat{\beta}_{IV}) \geq Var(\hat{\beta}_{OLS})$. If the regressors are correlated with the error, $\hat{\beta}_{OLS}$ is inconsistent while $\hat{\beta}_{IV}$ is consistent. If the regressors are uncorrelated with the error, $\hat{\beta}_{OLS}$ is more efficient than $\hat{\beta}_{IV}$ (which remains consistent).
- Therefore, we would like to test the null hypothesis of no correlation between the regressors and the error term. Hausman propose the following test statistic:

$$H = (\hat{\beta}_{IV} - \hat{\beta}_{OLS})'(Var(\hat{\beta}_{IV}) - Var(\hat{\beta}_{OLS}))^{-1}(\hat{\beta}_{IV} - \hat{\beta}_{OLS}) \underset{H_0}{\overset{a}{\sim}} \chi^2(k).$$

Model Mis-Specification

Examples using simulated data

- In Chapter 1 we considered the baseline linear regression models estimated by OLS.
  - We know that OLS estimators are BLUE (best linear unbiased estimator) if the conditions of the Gauss-Markov theorem (GMT) are satisfied. The following are the conditions: Linearity, Homoskedasticity and lack of correlation, No collinearity.
  - In addition, the errors are required to be normally distributed for the OLS estimators to have a finite sample normal distribution, which also justifies the finite sample distribution of the $t-$ and $F-$statistics.
  - Furthermore, model parameters are implicitly assumed to be stable over time
  - The regressors, when stochastic, are assumed to be uncorrelated with the error term.

- **Homoskedasticity** To test the null hypothesis of homoskedasticity, we perform two tests: the Breusch-Pagan-Godfrey (BPGT) test and the White test (WT), described in the previous sections. The outcome, described in Table 1, leads to rejection of the null hypothesis for the BPGT test.

F-statistic	7.577	Prob.F(1,198)		0.006
Obs*R-squared	7.372	Prob. Chi-Sq(1)		0.007
Scaled explained SS	5.715	Prob. Chi-Sq(1)		0.017
Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	82.648	7.625	10.839	0.000
X	3.669	1.333	2.753	0.006
R-squared	0.037	Mean dep var		85.882

Table 1: *Breusch-Pagan-Godfrey homoskedasticity test*

- Table 2 also presents the output for the WT test and reject the null hypothesis.

F-statistic	4.304	Prob.F(2,197)		0.015
Obs*R-squared	8.373	Prob. Chi-Sq(2)		0.015
Scaled explained SS	6.491	Prob. Chi-Sq(2)		0.039
Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	73.928	11.489	6.435	0.000
X^2	0.270	0.266	1.015	0.312
X	3.537	1.339	2.642	0.009
R-squared	0.042	Mean dep var		85.882

Table 2: *White heteroskedasticity test*

- **No Serial Correlation**
  - To check if the residuals are not serially correlated. We perform the Durbin-Watson (DW) statistic and the Breusch-Godfrey Lagrange Multipliers (LM) test.
  - The null hypothesis of the Durbin-Watson test is no serial correlation, the alternative is first order correlation, i.e., $H_1 : \varepsilon_t = \rho\varepsilon_{t-1} + u_t, \ \rho \neq 0$. The DW statistic is equal to 0.9545, much lower than the theoretical value under the null hypothesis, which is 2.
  - This provides informal support against the null hypothesis, meaning that the evidence suggests the errors are autocorrelated.

Examples using simulated data

- The LM test permits one to choose the serial correlation order under the alternative hypothesis. The results of the LM test reported in Table 3.

F-statistic	18.882	Prob. F(4,194)		0.000
Obs*R-squared	56.045	Prob. Chi-Sq(4)		0.000
Variable	Coefficient	Std.Error	t-Statistic	Prob.
ALPHA(1)	-0.003	0.572	-0.005	0.996
ALPHA(2)	-0.032	0.101	-0.314	0.754
RESID(-1)	0.492	0.072	6.807	0.000
RESID(-2)	0.097	0.080	1.213	0.227
RESID(-3)	-0.025	0.081	-0.306	0.760
RESID(-4)	-0.087	0.073	-1.190	0.235
R-squared	0.280	Mean dep var		-0.000

Table 3: *Breusch-Godfrey LM test*

- **Normality** Descriptive information about normality of the error term is provided by the histogram of the residuals. The resulting graph is displayed in Figure 1.

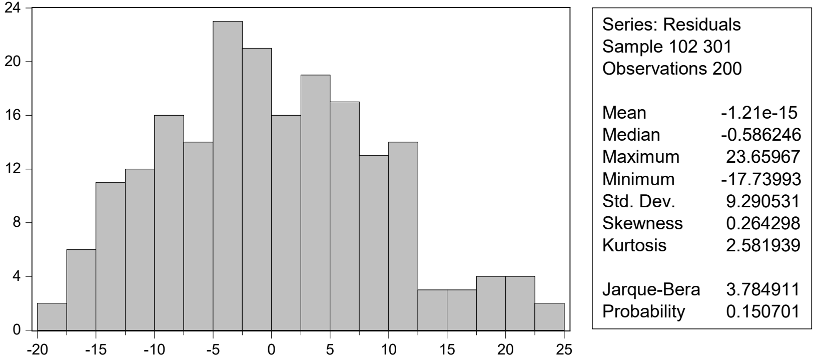


Figure 1: *Normality test: Histogram*

Examples using simulated data

- **Parameter Stability** To test whether the model parameters are constant for the whole estimation sample against the alternative of a break in a specific date,
  - we can apply the Chow breakpoint or the Chow forecast tests, depending on whether or not the break date is such that we can re-estimate the model in both the pre- and the post-break samples (see the theory part for more details).
  - we can compute a set of recursive statistics, based on the recursive residuals, the one-step ahead forecast errors, and the recursive coefficients. These are particularly useful when the break date is unknown.

Examples using simulated data

- We start with the Chow breakpoint test and consider three possible alternative breakpoints, corresponding to t = 151 or t = 201 or t = 251.
- All the three test outcomes, reported in Table 4, lead to rejection of the null hypothesis at the 5% confidence level.

Examples using simulated data

Chow Breakpoint Test: 151			
Equation Sample: 102 301			
F-statistic	12.536	Prob.F(2,196)	0.000
Log likelihood ratio	24.074	Prob. Chi-Sq(2)	0.000
Wald Statistic	25.071	Prob. Chi-Sq(2)	0.000
Chow Breakpoint Test: 201			
Equation Sample: 102 301			
F-statistic	27.779	Prob.F(2,196)	0.000
Log likelihood ratio	49.912	Prob. Chi-Sq(2)	0.000
Wald Statistic	55.558	Prob. Chi-Sq(2)	0.000
Chow Breakpoint Test:251			
Equation Sample: 102 301			
F-statistic	14.526	Prob.F(2,196)	0.000
Log likelihood ratio	27.643	Prob.Chi-Sq(2)	0.000
Wald Statistic	29.051	Prob. Chi-Sq(2)	0.000

Table 4: *Chow breakpoint tests*

Examples using simulated data

- Assume that we want to test for a break at observation 301 (so that we have only one observation in the second subsample and we are in the case $T_2 < k$). The outcome is reported in Table 5. The LR test statistic is asymptotically χ^2 distributed with degrees of freedom equal to the number of dummies. It turns out that stability at the end of the estimation sample (observation 301) is rejected at the 10% significance level.

Test predictions for observations from 301 to 301			
	Value	df	Probability
F-statistic	3.714	(1, 197)	0.055
Likelihood ratio	3.735	1	0.053

Table 5: *Chow forecast test*

Examples using simulated data

- If the breakpoint is now known in advance, we can implement Bai-Perron sequential test. The results are reported in Table 6.
- The Bai-Perron test detects breaks in the intercept for observations 146, 207, 237, 268. At the same time the null hypothesis of no breaks in X is rejected for observations 146, 176, 207, 237, 268.
- In this case we know the DGP and the fact that there is a single break in observation 201, the Bai and Perron test is incorrect. Likely due to short sample available and size of the break.

Examples using simulated data

Variable	Coefficient	Std. Error	t-Statistic	Prob.
102-145 – 44 obs				
C	4.523	1.212	3.733	0.000
X	0.701	0.193	3.622	0.000
146-175 – 30 obs				
C	0.841	1.452	0.579	0.563
X	0.750	0.256	2.931	0.004
176-206 – 31 obs				
C	3.990	1.462	2.730	0.007
X	1.368	0.264	5.185	0.000
207-236 – 30 obs				
C	9.152	1.516	6.038	0.000
X	1.800	0.277	6.496	0.000
237-267 – 31 obs				
C	10.900	1.558	6.999	0.000
X	1.926	0.291	6.616	0.000
268-301 – 34 obs				
C	3.302	1.360	2.428	0.016
X	2.360	0.237	9.946	0.000

Table 6: *Bai and Perron breakpoint test*

Examples using simulated data

- The graphs of the recursively computed OLS estimates of the model parameters are displayed in Figure 2 and Figure 3. In the presence of parameter instability, there should be significant time variation in the plots.

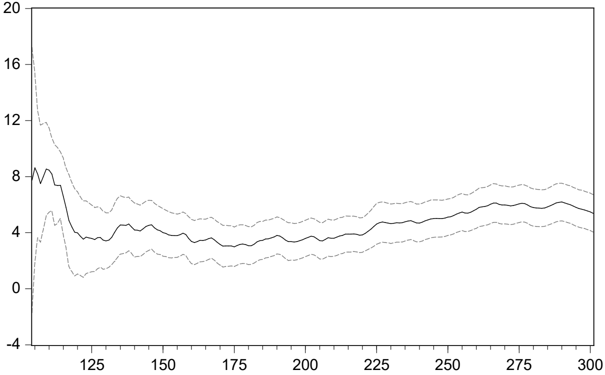


Figure 2: *Recursive* α_1 *estimates* $\pm$ *2 s.e.*

Examples using simulated data

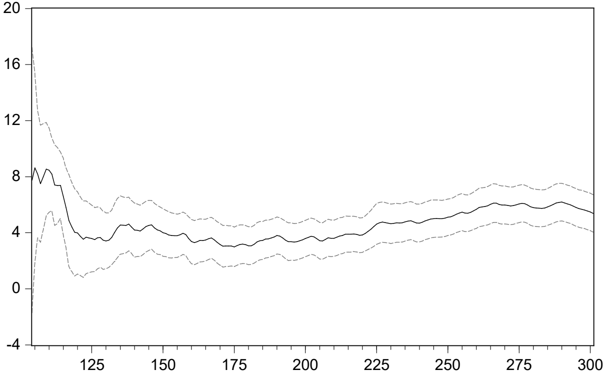


Figure 3: *Recursive α_2 estimates $\pm$ 2 s.e.*

- **Dummy variables**
  From the DGP, we know that the model parameters indeed change after t = 200. We can define a dummy variable whose value is 0 when $t <$ 201 and 1 afterwards. We then add the dummy to the starting model, which becomes:

$$y = \alpha_0 + \beta_1 D + \alpha_1 x + \beta_2 Dx + u,$$

  Estimation results, presented in Table 7, indicate that both β_1 and β_2 are statistically significant at the 5% level.

	Coefficient	Std. Error	t-Statistic	Prob.
ALPHA(0)	3.673	0.837	4.390	0.000
BETA(1)	3.221	1.187	2.714	0.007
ALPHA(1)	0.884	0.141	6.279	0.000
BETA(2)	1.316	0.208	6.320	0.000
R-squared	0.571	Mean dep var		6.695
Adjusted R-squared	0.564	S.D. dep var		12.560
S.E. of regression	8.290	Akaike IC		7.088
Sum squared resid	13468.600	Schwarz IC		7.154
Log likelihood	-704.768	Hannan-Quinn		7.114
F-statistic	86.940	DW stat		1.029
Prob(F-statistic)	0.000			

Table 7: *Model with dummy variable*

Examples using simulated data

- Since we know the DGP, we can also check if the estimated coefficients are significantly different from the true parameter values. The results are as follows:
  - For $\alpha(0)$, $t - stat$ = $\frac{3.6728-1}{0.8366} = 3.19$, for $\alpha(1)$, $t - stat$ = $\frac{0.8841-1}{0.1408} = -0.82$
  - For $\beta(1)$, $t - stat = \frac{3.2213-1}{1.1869} = 1.87$, for $\beta(2)$, $t - stat = \frac{1.3156-1}{0.2082} = 1.52$
- Hence, at the 5% confidence level, the last three parameters are not different from the true values, but the first one is.

- The final step is the forecast computation and evaluation. We use 102 - 301 as estimation sample and 302 - 501 as forecast sample.
- We want to assess the effects of ignoring the parameter break on the model forecasting performance. The forecasts from the models with and without the dummies are presented in Figure 4, while Table 8 reports the RMSFE and MAFE.

	Forecasting Method	
	Static	*Recursive*
RMSFE	10.1278	9.6421
MAFE	8.3141	7.9382

Table 8: *Forecasting performance model with dummy variable*

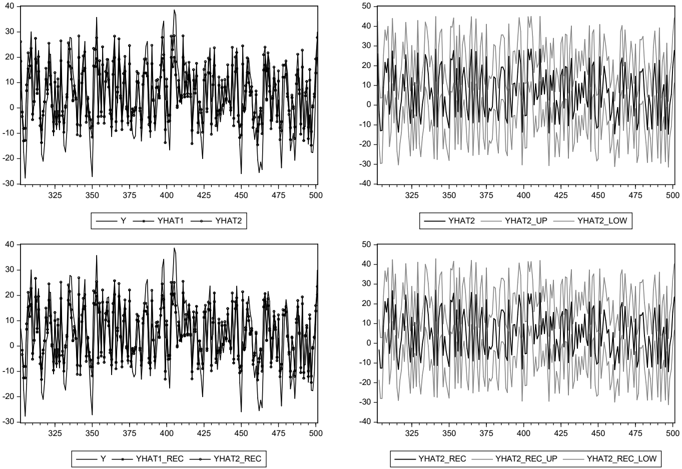


Figure 4: *Dummy regression: Simple and recursive forecasts vs forecast of the simple OLS regression*

Model Mis-Specification

- **Homoskedasticity**
  - Let us focus on the Model 2 formulation for Euro area GDP growth that, from Chapter 1, is:

    $$y_t = \beta_0 + \beta_1 ipr_t + \beta_2 su_t + \beta_3 sr_t + \varepsilon_t,$$

    and consider the full sample period 1996Q1 to 2013Q2.
  - Similary, to test the null hypothesis of homoskedasticity, we perform two tests: the Breusch-Pagan-Godfrey (BPGT) test and the White test (WT), described in the previous sections.

- The Breusch-Pagan-Godfrey test output is shown in Table 9 .

F-statistic	0.798	Prob. F(3,66)		0.499
Obs*R-squared	2.451	Prob. Chi-Sq(3)		0.484
Scaled explained SS	1.930	Prob.Chi-Sq(3)		0.587
Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	0.075	0.013	5.617	0.000
IPR	0.009	0.010	0.851	0.398
SU	-0.005	0.005	-0.976	0.333
SR	0.002	0.002	1.283	0.204
R-squared	0.035	Mean dep var		0.080

Table 9: *Breusch-Pagan-Godfrey test for heteroskedasticity.*

- The null hypothesis is $H_0 : \sigma_i^2 = \sigma^2$, i.e., homoskedasticity, against the alternative that there is an unknown relationship between the error variance and one or a set of regressors (or function thereof).
- Under the null hypothesis, the statistic follows a χ^2 distribution with as many degrees of freedom of 3. The null hypothesis cannot be rejected at 1%, 5%, and 10% significance levels, indicating no heteroskedasticity in the residuals.

- Table 10 presents the output of the White test. The null hypothesis of the White test is the same as that of the Breusch-Pagan-Godfrey. The null hypothesis of homoskedasticity cannot be rejected at conventional significance levels.

F-statistic	0.510	Prob.F(9,60)		0.861
Obs*R-squared	4.977	Prob. Chi-Sq(9)		0.836
Scaled explained SS	3.920	Prob. Chi-Sq(9)		0.917
Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	0.058	0.022	2.710	0.009
IPR	0.009	0.018	0.486	0.629
IPR2	0.000	0.006	0.021	0.983
IPR*SU	-0.001	0.010	-0.120	0.905
IPR*SR	0.000	0.004	0.086	0.932
SU	-0.004	0.008	-0.471	0.639
SU2	0.001	0.002	0.696	0.489
SU*SR	-0.001	0.001	-0.667	0.507
SR	0.002	0.003	0.708	0.482
SR2	0.000	0.000	1.103	0.275
R-squared	0.071	Mean dep var		0.080

Table 10: *White test for heteroskedasticity.*

- **No Serial Correlation**
  - To check if the residuals are not serially correlated. We perfom the Durbin-Watson (DW) statistic and the Breusch-Godfrey Lagrange Multipliers (LM) test.
  - The null hypothesis of the Durbin-Watson test is no correlation, against the alternative of first order correlation, i.e., $H_1 : \varepsilon_t = \rho\varepsilon_{t-1} + u_t, \ \rho \neq 0$. The DW test statistic is 1.433661.
  - While the DW test is for testing serial correlation of first order under the alternative, higher order correlation can be allowed by the LM test.

Forecasting Euro area GDP growth

- Tables 11 and 12 present the LM test with correlation of orders 2 and 3.

F-statistic	3.664	Prob.F(2,64)		0.031
Obs*R-squared	7.191	Prob. Chi-Sq(2)		0.028
Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	0.002	0.035	0.051	0.960
IPR	-0.015	0.028	-0.520	0.605
SU	0.005	0.014	0.333	0.740
SR	-0.002	0.005	-0.346	0.731
RESID(-1)	0.230	0.124	1.858	0.068
RESID(-2)	0.199	0.133	1.493	0.140
R-squared	0.103	Mean dep var		0.000

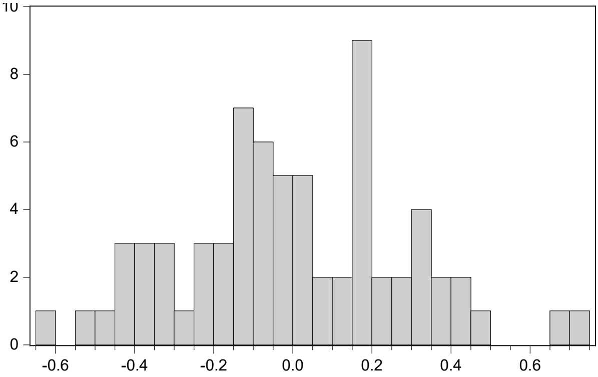
Table 11: *Serial correlation LM test with $m = 2$*

- The null hypothesis is rejected by both tests at the 5% significance level, indicating that the residuals are serially correlated.

F-statistic	4.439	Prob.F(3,63)		0.007
Obs*R-squared	12.214	Prob. Chi-Sq(3)		0.007
Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	0.003	0.034	0.090	0.929
IPR	-0.029	0.028	-1.035	0.305
SU	0.013	0.014	0.883	0.381
SR	-0.003	0.005	-0.612	0.542
RESID(-1)	0.180	0.122	1.476	0.145
RESID(-2)	0.165	0.130	1.267	0.210
RESID(-3)	0.301	0.128	2.340	0.022
R-squared	0.174	Mean dep var		0.000

Table 12: *Breusch-Godfrey serial correlation LM test with $m = 3$*

- **Normality** Figure 5 reports the required histogram and a set of statistics on the residuals of Model 2. The null hypothesis of normality cannot be rejected at conventional significance levels.



Series: Residuals
Sample 1996Q1 2013Q2
Observations 70

Mean	2.22e-17
Median	-0.018580
Maximum	0.722292
Minimum	-0.633967
Std. Dev.	0.284140
Skewness	0.144000
Kurtosis	2.771743
Jarque-Bera	0.393884
Probability	0.821238

Figure 5: *Histogram of the errors in Model 2*

- **Parameter stability**
  - We illustrate the use of Chow breakpoint test by investigating whether parameter instability or structural breaks exist in two occasions within our full sample period: the early 2000s economic recession that affected the European Union mostly around 2001 - 2002, and the recent 2007/2008 financial crisis.
  - In principle, we should first test that the error variance does not change over the two subperiods, but since we did not reject homoskedasticity we take the variance to be constant.

F-statistic	2.045	Prob. F(4,62)	0.099
Log likelihood ratio	8.674	Prob. Chi-Sq(4)	0.070
Wald Statistic	8.179	Prob. Chi-Sq(4)	0.085

Table 13: *Chow breakpoint test: Breakpoint in 2001Q1*

F-statistic	3.513	Prob. F(4,62)	0.012
Log likelihood ratio	14.301	Prob. Chi-Sq(4)	0.006
Wald Statistic	14.053	Prob. Chi-Sq(4)	0.007

Table 14: *Chow breakpoint test: Breakpoint in 2001Q2*

F-statistic	7.617	Prob. F(4,62)	0.000
Log likelihood ratio	27.980	Prob. Chi-Sq(4)	0.000
Wald Statistic	30.467	Prob. Chi-Sq(4)	0.000

Table 15: *Chow breakpoint test: Breakpoint in 2007Q3*

- Tables 13 and 14 present the two tests for whether there is a structural change in the parameters in 2001. At the 5% significance level, the null hypothesis of no break in 2001Q1 cannot be rejected, but the hypothesis of no break at 2001Q2 can be rejected. Table 15 presents the same test but with 2007Q3 as the break date. The outcome shows a clear rejection of no breakpoint also at that time.

- Suppose we are interested in whether there is a breakpoint in 2012Q3, close to the end of our full sample. The outcome of the tests, shown in Table 16, leads to rejection of the null of no break in 2012Q3 at the 5% level.

	Value	df	Probability
F-statistic	3.115	(4,62)	0.021
Likelihood ratio	12.819	4.000	0.012

Table 16: *Chow Forecast Test: Breakpoint in 2012Q3.*

- The major criticism of Chow-type tests is that prior knowledge about the break dates is required. However, there is also a need for tests that allow for the detection of structural changes without a pre-specified break date.
- Bai and Perron (1998) describe procedures for identifying the multiple breaks. When we implement the Bai-Perron sequential test to determine the optimal breaks in Model 2, we find the results reported in Table 17 and 18.

Forecasting Euro area GDP growth

- The Bai-Perron test detects breaks in 1998Q4, 2001Q2, 2004Q3, 2008Q2, and 2011Q1.

Variable	Coefficient	Std.Error	t-Statistic	Prob.
1996Q1 - 1998Q3 – 11obs				
C	0.499	0.104	4.777	0.000
IPR	0.431	0.147	2.930	0.005
SU	-0.037	0.048	-0.759	0.452
SR	-0.023	0.010	-2.371	0.022
1998Q4 - 2001Q1 – 10obs				
C	0.840	0.135	6.212	0.000
IPR	-0.045	0.131	-0.344	0.732
SU	0.082	0.051	1.602	0.116
SR	0.008	0.011	0.743	0.461
2001Q2 - 2004Q2 – 13obs				
C	0.280	0.063	4.461	0.000
IPR	0.251	0.126	1.998	0.052
SU	-0.027	0.034	-0.773	0.444
SR	0.011	0.009	1.307	0.198

Table 17: *Estimation output of Model 2 with breaks detected with the Bai and Perron breakpoint test 1.*

Forecasting Euro area GDP growth

Variable	Coefficient	Std.Error	t-Statistic	Prob.
2004Q3 - 2008Q1 – 15obs				
C	0.341	0.122	2.793	0.008
IPR	0.390	0.142	2.745	0.009
SU	0.022	0.037	0.599	0.552
SR	0.002	0.011	0.222	0.825
2008Q2 - 2010Q4 – 11obs				
C	-0.051	0.094	-0.544	0.589
IPR	0.273	0.034	7.992	0.000
SU	0.016	0.034	0.462	0.646
SR	-0.015	0.017	-0.893	0.376
2011Q1 - 2013Q2 – 10obs				
C	0.044	0.081	0.537	0.594
IPR	0.280	0.079	3.541	0.001
SU	0.005	0.030	0.162	0.872
SR	0.000	0.012	0.036	0.972

Table 18: *Estimation output of Model 2 with breaks detected with the Bai and Perron breakpoint test 2.*

- As an alternative, we plot the recursive residuals in Figure 6, together with the plus and minus 2 times standard error bands. Evidence of parameter instability is found when the recursive residuals are outside the bands.

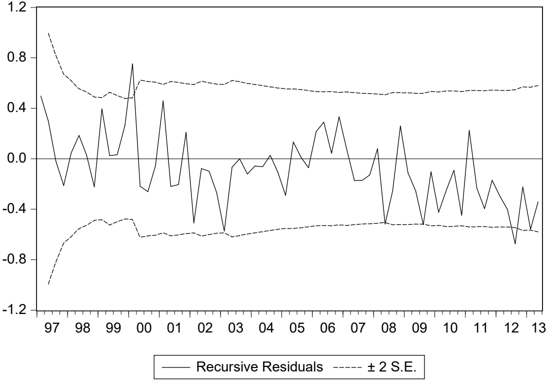


Figure 6: *Plot of recursive with the* $\pm 2 \times$ *s.e. confidence bounds.*

- Next we compute the one-step forecast test, based on the plot in Figure 7, and assuming a 5% significance level, evidence of parameter instability can be detected in 2000Q1, 2008Q2, 2009Q3, and 2012Q3.

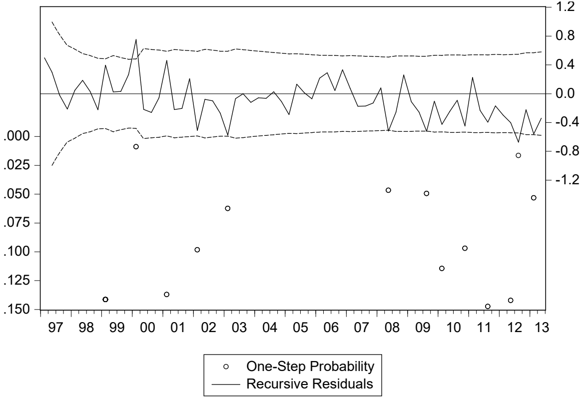


Figure 7: *One-step forecast test*

- The recursively estimated coefficients, labeled as C(1) to C(4) in Figure 8, corresponding to the coefficients of the constant term, IP growth, growth in the ESI, and stock returns.
- In the presence of parameter instability, there should be significant variation in the evolution of the estimators. We observe some marked changes prior to 2002 and some movement around 2008Q3 and 2008Q4.

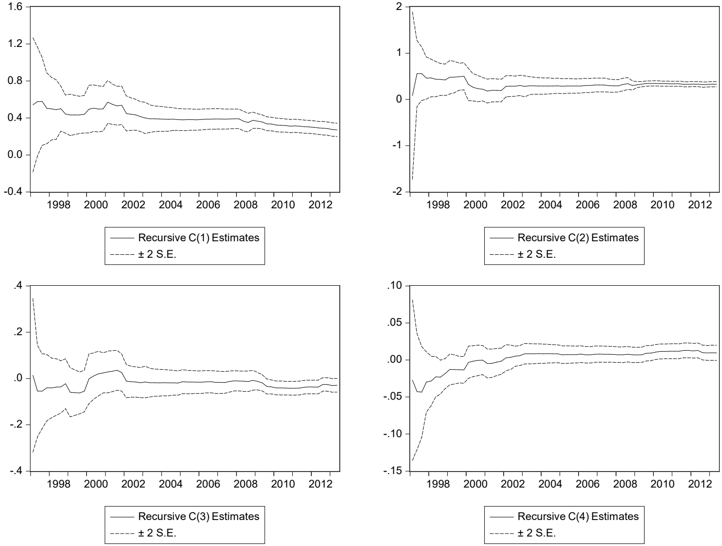


Figure 8: *Plots of recursive coefficients alongside the* $\pm 2 \times$ *s.e. confidence bounds*

- **Dummy variables** The parameter stability analysis provided evidence for some possible breaks since the second half of 2007, in coincidence with the financial crisis and the resulting the Euro area recession, and around 2000 - 2001, soon after the introduction of the euro and corresponding to the early 2000s recession. We therefore create two dummy variables. The first, called $d_t^{EA\_cri}$, has the value of 1 for the period 2007Q2 to 2013Q1, and 0 otherwise. The second, called d_t^{2000s}, has the value of 1 for the period 2000Q1 to 2001Q4, and 0 otherwise.

- Model 2.1 contains the independent variables in Model 2 plus the dummy for the Euro area crisis (Table 19).

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.375	0.040	9.271	0.000
IPR	0.301	0.025	11.812	0.000
SU	-0.021	0.013	-1.568	0.122
SR	0.006	0.005	1.196	0.236
Dea	-0.302	0.071	-4.278	0.000
R-squared	0.841	Mean dep var		0.350
Adjusted R-squared	0.831	S.D. dep var		0.629
S.E. of regression	0.259	Akaike IC		0.202
Sum squared resid	4.347	Schwarz IC		0.362
Log likelihood	-2.060	Hannan-Quinn		0.266
F-statistic	85.767	DW stat		1.792
Prob(F-statistic)	0.000			

Table 19: *Estimation output of Model 2.1*

Forecasting Euro area GDP growth

- Model 2.2 corresponds to Model 2 plus the dummy for the early 2000s recession (Table 20).

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.247	0.037	6.705	0.000
IPR	0.314	0.028	11.302	0.000
SU	-0.022	0.015	-1.487	0.142
SR	0.009	0.005	1.906	0.061
D2000s	0.224	0.111	2.016	0.048
R-squared	0.808	Mean dep var		0.350
Adjusted R-squared	0.796	S.D. dep var		0.629
S.E. of regression	0.284	Akaike IC		0.389
Sum squared resid	5.243	Schwarz IC		0.550
Log likelihood	-8.619	Hannan-Quinn		0.453
F-statistic	68.334	DW stat		1.462
Prob(F-statistic)	0.000			

Table 20: *Estimation output of Model 2.2*

Forecasting Euro area GDP growth

- The mathematical presentations of Models 2.1 and 2.2 are as follows:

$$Model\ 2.1:\ y_t = \alpha_0 + \beta_0 ipr_t + \gamma_0 su_t + \lambda_0 sr_t + \eta_0 d_t^{EA\_cri} + \varepsilon_t$$

$$Model\ 2.2:\ y_t = \alpha_0 + \beta_0 ipr_t + \gamma_0 su_t + \lambda_0 sr_t + \eta_1 d_t^{2000s} + \varepsilon_t$$

Our final forecasting specification, Model 2.3 (Table 21), is Model 2 with the dummies for both the financial crisis and the early 2000s recession:

$$Model\ 2.3:\ y_t = \alpha_0 + \beta_0 ipr_t + \gamma_0 su_t + \lambda_0 sr_t + \eta_0 d_t^{EA\_cri} + \eta_1 d_t^{2000s} + \varepsilon_t$$

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.381	0.043	8.883	0.000
IPR	0.302	0.027	11.156	0.000
SU	-0.023	0.015	-1.536	0.130
SR	0.006	0.005	1.349	0.183
Dea	-0.291	0.084	-3.473	0.001
D2000s	0.085	0.100	0.847	0.401
R-squared	0.868	Mean dep var		0.414
Adjusted R-squared	0.855	S.D. dep var		0.644
S.E. of regression	0.245	Akaike IC		0.119
Sum squared resid	3.238	Schwarz IC		0.328
Log likelihood	2.442	Hannan-Quinn		0.201
F-statistic	70.847	DW stat		1.965
Prob(F-statistic)	0.000			

Table 21: *Estimation output of Model 2.3*

- To highlight the impact on forecast accuracy of the dummies we use here 1996Q1 to 2010Q4 as estimation sample, leaving the last 10 quarters between 2011Q1 and 2013Q2 as the new forecast evaluation period. Next, we perform a forecasting exercise using these two models for the newly defined forecast evaluation period, with results shown in Figure 9.

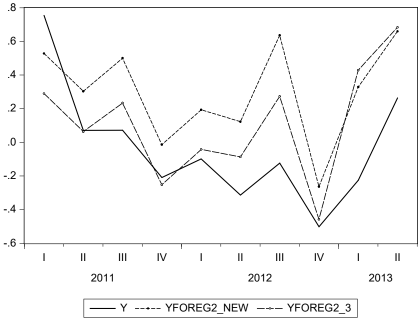


Figure 9: *Plot of forecasted GDP growth produced by Model 2 and by Model 2.3 alongside the actual GDP growth.*

- To further confirm this finding, we compute the RMSFE and the MAFE. Table 22 compares the forecast evaluation statistics, confirming the superiority of Model 2.3 and the importance of allowing for parameter change when forecasting.

	Model 2	Model 2 with dummies
RMSFE	0.412	0.326
MAFE	0.375	0.247

Table 22: *Forecast evaluation statistics: Model 2 without and with dummies (Model 2.3)*

Forecasting US GDP growth

- We used the US GDP data and focusing on Model 2, which is:

$$y_t = \beta_0 + \beta_1 ipr_t + \beta_2 su_t + \beta_3 sr_t + \varepsilon_t.$$

The full sample covers the period from 1985Q1 to 2013Q4. The estimation results are shown in Table 23.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.469	0.044	10.713	0.000
IPR	0.312	0.034	9.195	0.000
SU	0.007	0.007	0.979	0.330
SR	0.013	0.007	1.833	0.070
R-squared	0.523	Mean dep var		0.657
Adjusted R-squared	0.510	S.D. dep var		0.600
S.E. of regression	0.420	Akaike IC		1.137
Sum squared resid	19.763	Schwarz IC		1.232
Log likelihood	-61.951	Hannan-Quinn		1.176
F-statistic	40.919	DW stat		2.033
Prob(F-statistic)	0.000			

Table 23: *Estimation output for Model 2*

- **Homoskedasticity**
  The Breusch-Pagan-Godfrey test output is shown in Table 24. The null hypothesis of homoskedasticity cannot be rejected.

F-statistic	0.572	Prob. F(3,112)		0.635
Obs*R-squared	1.751	Prob. Chi-Sq(3)		0.626
Scaled explained SS	1.593	Prob. Chi-Sq(3)		0.661
Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	0.176	0.025	7.023	0.000
IPR	0.003	0.019	0.158	0.875
SU	-0.001	0.004	-0.369	0.713
SR	-0.004	0.004	-0.931	0.354
R-squared	0.015	Mean dep var		0.170

Table 24: *Breusch-Pagan-Godfrey heteroskedasticity test*

Forecasting US GDP growth

- While the output of the White test is presented in Table 25.

F-statistic	0.697268	Prob. F(9,106)		0.7100
Obs*R-squared	6.484	Prob. Chi-Sq(9)		0.691
Scaled explained SS	5.901	Prob. Chi-Sq(9)		0.750
Variable	Coefficient	Std.Error	t-Statistic	Prob.
C	0.172	0.039	4.395	0.000
IPR^2	0.003	0.011	0.266	0.791
$IPR * SU$	0.004	0.003	1.200	0.233
$IPR * SR$	-0.003	0.003	-1.017	0.311
IPR	-0.009	0.024	-0.380	0.705
SU^2	-0.000	0.000	-1.044	0.299
$SU * SR$	0.001	0.001	0.889	0.376
SU	-0.005	0.005	-0.927	0.356
SR^2	0.000	0.000	0.262	0.793
SR	0.001	0.005	0.201	0.841
R-squared	0.056	Mean dep var		0.170

Table 25: *White heteroskedasticity test*

- **No Serial Correlation** The diagnostic tests we consider here are the Durbin-Watson (DW) test and the Breusch-Godfrey LM test. The DW statistic is 2.033324, close to the theoretical value under the null of uncorrelated residuals. The LM tests confirm the same findings - we skip the details here.

- **Parameter stability** As a first step, we implement the Chow breakpoint test to examine whether there was a breakpoint in, respectively, 2001Q1, 2001Q2, 2007Q3, and 2007Q4. Table 26 presents the results for the CH(1) and CH(2) statistics.

Chow Breakpoint Test: 2001Q1			
$CH(1)$	3.264	Prob. F(4,108)	0.014
$CH(2)$	0.999	Prob. F(47,61)	0.496
Chow Breakpoint Test: 2001Q2			
$CH(1)$	3.139	Prob. F(4,108)	0.017
$CH(2)$	0.977	Prob. F(46,62)	0.528
Chow Breakpoint Test: 2007Q3			
$CH(1)$	4.025	Prob .F(4,108)	0.004
$CH(2)$	1.055	Prob. F(21,87)	0.410
Chow Breakpoint Test: 2007Q4			
$CH(1)$	4.314	Prob. F(4,108)	0.003
$CH(2)$	1.120	Prob. F(20,88)	0.345

Table 26: *Chow breakpoint test*

- To test for breaks without exogenously setting up the breakpoint data we use the Bai-Perron test. Results of the test are described in Table 27 and 28.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
1985Q1 - 1989Q3 – 19obs				
C	0.963	0.139	6.940	0.000
IPR	0.045	0.117	0.382	0.703
SU	-0.019	0.027	-0.677	0.500
SR	-0.017	0.015	-1.147	0.254
1989Q4 - 1996Q1 – 26obs				
C	0.328	0.122	2.680	0.009
IPR	0.416	0.094	4.441	0.000
SU	0.012	0.012	0.989	0.325
SR	0.002	0.022	0.072	0.943
1996Q2 - 2000Q2 – 17obs				
C	1.005	0.251	4.001	0.000
IPR	0.185	0.165	1.121	0.265

Table 27: *Bai-Perron breakpoint test1*

Variable	Coefficient	Std. Error	t-Statistic	Prob.
SU	-0.076	0.040	-1.881	0.063
SR	-0.007	0.026	-0.274	0.785
2000Q3 - 2004Q3 – 17obs				
C	0.604	0.088	6.832	0.000
IPR	0.248	0.100	2.487	0.015
SU	0.017	0.022	0.757	0.451
SR	0.023	0.019	1.180	0.241
2004Q4 - 2008Q4 – 17obs				
C	0.397	0.094	4.227	0.000
IPR	0.045	0.099	0.456	0.650
SU	0.001	0.012	0.053	0.958
SR	0.077	0.018	4.327	0.000
2009Q1 - 2013Q4 – 20obs				
C	0.286	0.092	3.118	0.002
IPR	0.251	0.057	4.430	0.000
SU	0.004	0.014	0.283	0.778
SR	0.008	0.019	0.403	0.688

Table 28: *Bai-Perron breakpoint test 2*

- Suppose we are interested in whether there is a breakpoint in 2012Q3, close to the end of our full sample. The results for the Chow forecast tests, $CH(3)$ and $CH(4)$ appearing in equations (28) and (29), are respectively 0.704589 (p-value 0.6465) and 4.536486 (p=value 0.6045). Both of them suggest no rejection of the null of no break in 2012Q3, at the 5% significance level.

- We can also examine a recursive residual plot and compute the one-step forecast test, as shown in Figure 10. Evidence of parameter instability can be detected in the beginning of the 1990s, 2000Q1, 2004Q1, 2007Q2, and 2010Q3.

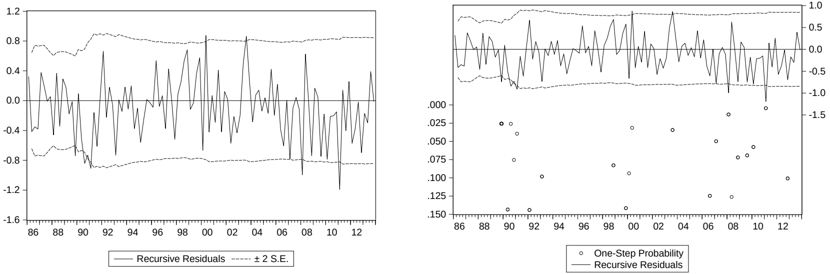


Figure 10: *Recursive estimation Model 2 for US GDP growth 1*

- Figure 11 produces recursive coefficients. For all four coefficients we observe significant variation before 1990, with a rather stable behavior afterwards.

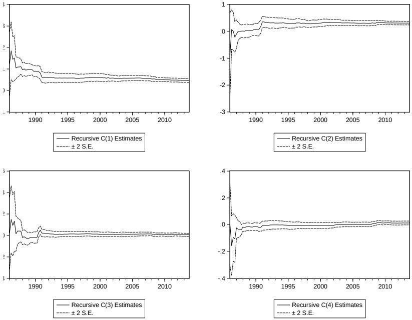


Figure 11: *Recursive estimation Model 2 for US GDP growth 2*

- **Dummy variables** We introduce dummy variables corresponding to the recent financial crisis (2007Q4 - 2009Q2), as well the early 2000s recession (2000Q1 - 2001Q4). We re-estimate Model 2 by adding these two dummies separately. Model 2.1 contains the independent variables in Model 2 plus the dummy for the financial crisis. Similarly, Model 2.2 consists of Model 2 plus the dummy for the early 2000s recession. They are:

$$Model\ 2.1:\ y_t = \alpha_0 + \beta_0 ipr_t + \gamma_0 su_t + \lambda_0 sr_t + \eta_0 d_t^{fincris} + \varepsilon_t$$

$$Model\ 2.2:\ y_t = \alpha_0 + \beta_0 ipr_t + \gamma_0 su_t + \lambda_0 sr_t + \eta_1 d_t^{2000s} + \varepsilon_t$$

The dummy variables are insignificant in both models. Implying that the other explanatory variables seem sufficient to capture the drop in GDP growth in the two recessions.

- In order to see the impact on forecast accuracy of the inclusion of the financial crisis dummy, we define the new estimation sample as 1985Q1 to 2010Q4, leaving the last 10 quarters between 2011Q1 and 2013Q2 as the new forecast evaluation period. Our forecasting model is simply Model 2 augmented with the dummies for the financial crisis, labeled as model 2.3 as in the previous example:

$$Model\ 2.3: \ y_t = \alpha_0 + \beta_0 ipr_t + \gamma_0 su_t + \lambda_0 sr_t + \eta_0 d_t^{EA\_cri} + \eta_1 d_t^{2000s} + \varepsilon_t$$

- The estimation output (not reported) of Model 2.3 reveals that both dummy variables are insignificant.
- Yet, comparing Model 2 (cf. Table 23) with 2.3, it turns out that the latter has a better in-sample fit than Model 2, namely model 2.3 has larger values of R^2 and $\overline{R}^2$, and smaller values of AIC and HQ criteria.

- The evaluation statistics for the one-step ahead recursive forecasts and simple forecasts are presented in Table 29.

	Model 2	Model 2.3
	Static Forecasts	
RMSFE	0.522	0.519
MAFE	0.422	0.419
	Recursive Forecasts	
RMSFE	0.531	0.528
MAFE	0.419	0.416

Table 29: *Forecast evaluation statistics from the simple forecast for Model 2 and for Model 2 with dummy variables and one-step ahead recursive forecasts of Model 2 and Model 2 with dummies*

- We revisit the regression models estimated in the previous chapter using the sample from Jan. 1998 to Dec. 2015. In particular we estimate the following:

$$
\begin{aligned}
OAS_{t+1} &= \beta_0 + \beta_1 VIX_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 SENT_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 PMI_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 sp500_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 VIX_t + \beta_2 SENT_t + \beta_3 PMI_t + \beta_4 sp500_t + \varepsilon_{t+1}
\end{aligned}
$$

Default risk

- The Durbin-Watson statistics are respectively DW_1 = 0.5660, DW_2 = 0.1219, DW_3 = 0.2390, DW_4 = 0.2634, and DW_5 = 0.7541. Hence, in all cases the null of no autocorrelation is rejected.
- when we look at the Breusch-Pagan-Godfrey test to check for the presence of heteroskedasticity we find: BPG_1 = 58.60, p-value $\approx$ 0, BPG_2 = 17.90, p-value = 0.00002, BPG_3 = 24.76, p-value $\approx$ 0, BPG_4 = 7.65, p-value = 0.006, BPG_5 = 37.06, p-value $\approx$ 0. This means that there is also evidence for the presence of heteroskedasticity.

- Next, we construct a dummy variable D_t equal to one during the period Jan. 2008 until Dec. 2009 and estimate the regressions using the full sample of data ending in 2015:

$$\begin{aligned}
OAS_{t+1} &= \beta_0 + \beta_1 VIX_t + \beta_2 D_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 SENT_t + \beta_2 D_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 PMI_t + \beta_2 D_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 sp500_t + \beta_2 D_t + \varepsilon_{t+1}
\end{aligned}$$

and

$$\begin{aligned}
OAS_{t+1} &= \beta_0 + \beta_1 VIX_t + \beta_2 D_t + \beta_3 VIX_t \times D_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 SENT_t + \beta_2 D_t + \beta_3 SENT_t \times D_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 PMI_t + \beta_2 D_t + \beta_3 PMI_t \times D_t + \varepsilon_{t+1} \\
OAS_{t+1} &= \beta_0 + \beta_1 sp500_t + \beta_2 D_t + \beta_3 sp500_t \times D_t + \varepsilon_{t+1}
\end{aligned}$$

The results appear in Table 30, Table 31, Table 32, and Table 33.

Models 1 - 4	Coefficient	t-Statistic	Prob.	R^2	Adjusted R^2
C	0.7544	2.586	0.01		
	(0.2918) [0.7355]	[1.026]	[0.31]		
VIX(-1)	0.233	17.082	0.00	0.73	0.73
	(0.0137) [0.0404]	[5.784]	[0.00]		
DUMMY	2.6355	7.493	0.00		
	(0.3517) [1.0737]	[2.455]	[0.01]		
C	8.2971	7.126	0.00		
	(1.1643) [2.0298]	[4.088]	[0.00]		
SENT(-1)	-0.0327	-2.505	0.01	0.39	0.38
	(0.0131) [0.0256]	[-1.281]	[0.20]		
DUMMY	4.6191	8.185	0.00		
	(0.5643) [2.9572]	[1.562]	[0.12]		

Table 30: *Default risk models augmented with dummies – Intercept. Square brackets are HAC estimator corrected standard errors and t-statistics using pre-whitening (1 lag) and Newey-West estimator (12 lags) 1*

Models 1 - 4	Coefficient	t-Statistic	Prob.	R^2	Adjusted R^2
C	21.2224	15.121	0.00		
	(1.4035) [3.3970]	[6.247]	[0.00]		
PMI(-1)	-0.2991	-11.315	0.00	0.61	0.60
	(0.0264) [0.0639]	[-4.678]	[0.00]		
DUMMY	3.3139	7.846	0.00		
	(0.4223) [1.1226]	[2.952]	[0.00]		
C	5.5208	36.435	0.00		
	(0.1515) [0.4189]	[13.179]	[0.00]		
SP500(-1)	-0.1808	-5.568	0.00	0.45	0.45
	(0.0320) [0.0609]	[-2.966]	[0.00]		
DUMMY	5.1032	11.285	0.00		
	(0.4522) [2.7140]	[1.880]	[0.06]		

Table 31: *Default risk models augmented with dummies – Intercept. Square brackets are HAC estimator corrected standard errors and t-statistics using pre-whitening (1 lag) and Newey-West estimator (12 lags) 2*

Default risk

Models 1 - 4	Coefficient	t-Statistic	Prob.	R^2	Adjusted R^2
C	1.6136	5.142	0.00		
	(0.3138) [0.6596]	[2.446]	[0.02]		
VIX(-1)	0.1904	12.725	0.00	0.77	0.77
	(0.0150) [0.0379]	[5.017]	[0.00]		
DUMMY	-1.9445	-2.192	0.03		
	(0.8870) [1.2098]	[-1.607]	[0.11]		
VIX(-1) $\times$ DUMMY	0.160	5.561	0.00		
	(0.0289) [0.0425]	[3.773]	[0.00]		
C	7.2609	6.497	0.00		
	(1.1175) [1.8078]	[4.017]	[0.00]		
SENT(-1)	-0.0210	-1.674	0.10	0.46	0.45
	(0.0125) [0.0213]	[-0.986]	[0.33]		
DUMMY	28.2095	6.144	0.00		
	(4.5911) [10.750]	[2.624]	[0.01]		
SENT(-1) $\times$ DUMMY	-0.3587	-5.173	0.00		
	(0.0693) [0.1468]	[-2.443]	[0.02]		

Table 32: *Default risk models augmented with dummies – Slope. Square brackets are HAC estimator corrected standard errors and t-statistics using pre-whitening (1 lag) and Newey-West estimator (12 lags) 1.*

Default risk

Models 1 - 4	Coefficient	t-Statistic	Prob.	R^2	Adjusted R^2
C	17.0234	11.028	0.00		
	(1.5436) [3.0382]	[5.603]	[0.00]		
PMI(-1)	-0.2197	-7.548	0	0.65	0.65
	(0.0291) [0.0549]	[-4.003]	[0.00]		
DUMMY	17.4956	6.435	0.00		
	(2.7187) [4.7411]	[3.690]	[0.00]		
PMI(-1) $\times$ DUMMY	-0.2966	-5.273	0.00		
	(0.0563) [0.0911]	[-3.255]	[0.00]		
C	5.4843	36.876	0.00		
	(0.1487) [0.5109]	[10.734]	[0.00]		
SP500(-1)	-0.1226	-3.393	0.00	0.48	0.47
	(0.0361) [0.0405]	[-3.024]	[0.00]		
DUMMY	4.9813	11.215	0.00		
	(0.4441) [1.8399]	[2.707]	[0.01]		
SP500(-1) $\times$ DUMMY	-0.2321	-3.218	0.00		
	(0.0721) [0.0927]	[-2.505]	[0.01]		

Table 33: *Default risk models augmented with dummies – Slope. Square brackets are HAC estimator corrected standard errors and t-statistics using pre-whitening (1 lag) and Newey-West estimator (12 lags) 2.*

Table results summary

- The parameter estimates associated with the dummy variables are highly significant.
- Using the Chow test and January 2008 as a potential break point, the null hypothesis of no structural break is also rejected for all model specifications.
- There are in fact two sets of standard errors and t-statistics reported in both tables, namely the OLS ones in curly brackets and those obtained with a HAC variance estimator.
- The standard errors and t-statistics reported in Tables 30 , 31, 32 and 33 are based on pre-whitening (1 lag) and Newey-West estimator (12 lags).

Table results summary

- With the modified statistics, we do observe some differences in Table 30 and 31. It appears that the sentiment index is no longer significant, nor is the dummy in Model 2. The latter is also the case with the Model 4 dummy. Looking at the results in Table 32 and 33 we see that the sentiment index does become significant during the financial crisis, i.e., the interaction dummy with the sentiment index remains significant with the HAC corrected standard errors, although the index itself is insignificant.

Concluding remarks

- Model mis-specification tests are a key ingredient in the formulation of forecasting models. In this chapter we reviewed a battery of tests that are commonly applied in a regression setting, together with possible remedies, such as the use of dummy variables. One has to keep in mind that many of the tests and remedies discussed in this chapter also apply to more complex models - such as those studied in later chapters of this book.