Dynamic regression models: Estimation & forecasting

EC 361–001

Prof. Santetti
Spring 2024

Materials

Required readings:

- `Hyndman & Athanasopoulos, ch. 10`
    - sections 10.1—10.3.

Motivation

Motivation

Last time, the **intuition** behind the use of ARIMA errors in time-series regression was introduced.

Now, our task is to employ the **estimation** and **forecasting** steps using this methodology.

we will do this through an **example**.

An example: Phillips curve

An example: Phillips curve

In **Macroeconomic theory**, **inflationary** pressures can arise due to and lower **unemployment**.

- This association is summarized by the *Phillips curve*.

One of its many variations is the so-called ***accelerationist*** Phillips curve, especially for *unstable* scenarios.

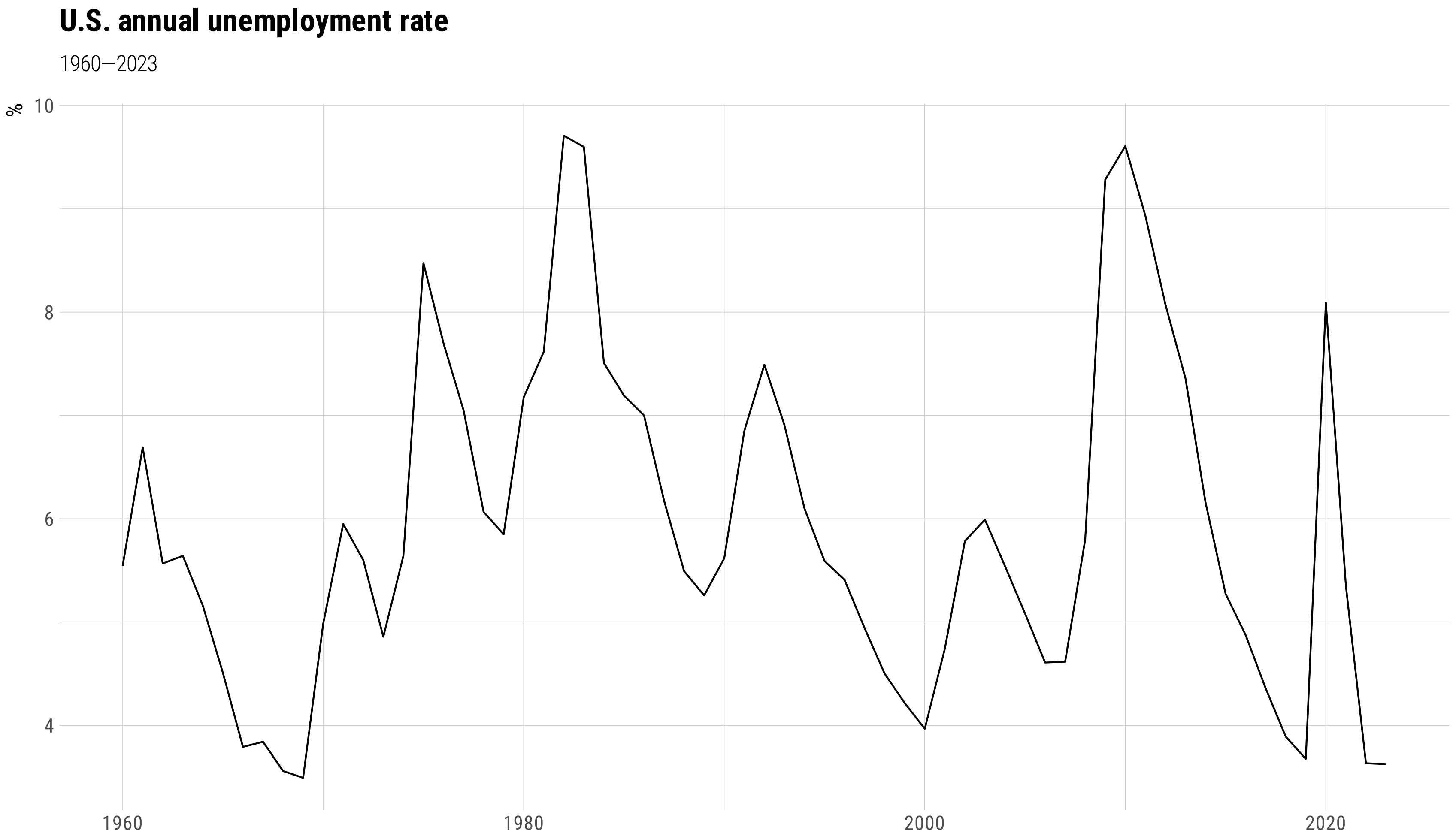
For statistical **modeling** purposes, the usual way to represent this relationship is the following:

$$\Delta\pi_t = \beta_0 + \beta_1 u_t + \varepsilon_t$$

where $\Delta\pi_t$ is the *change in the inflation rate*, and u_t is the *unemployment rate*.

It is expected that the **slope**, here represented by β_1, is **negative**, as higher employment is expected to generate *inflationary pressures*.

An example: Phillips curve

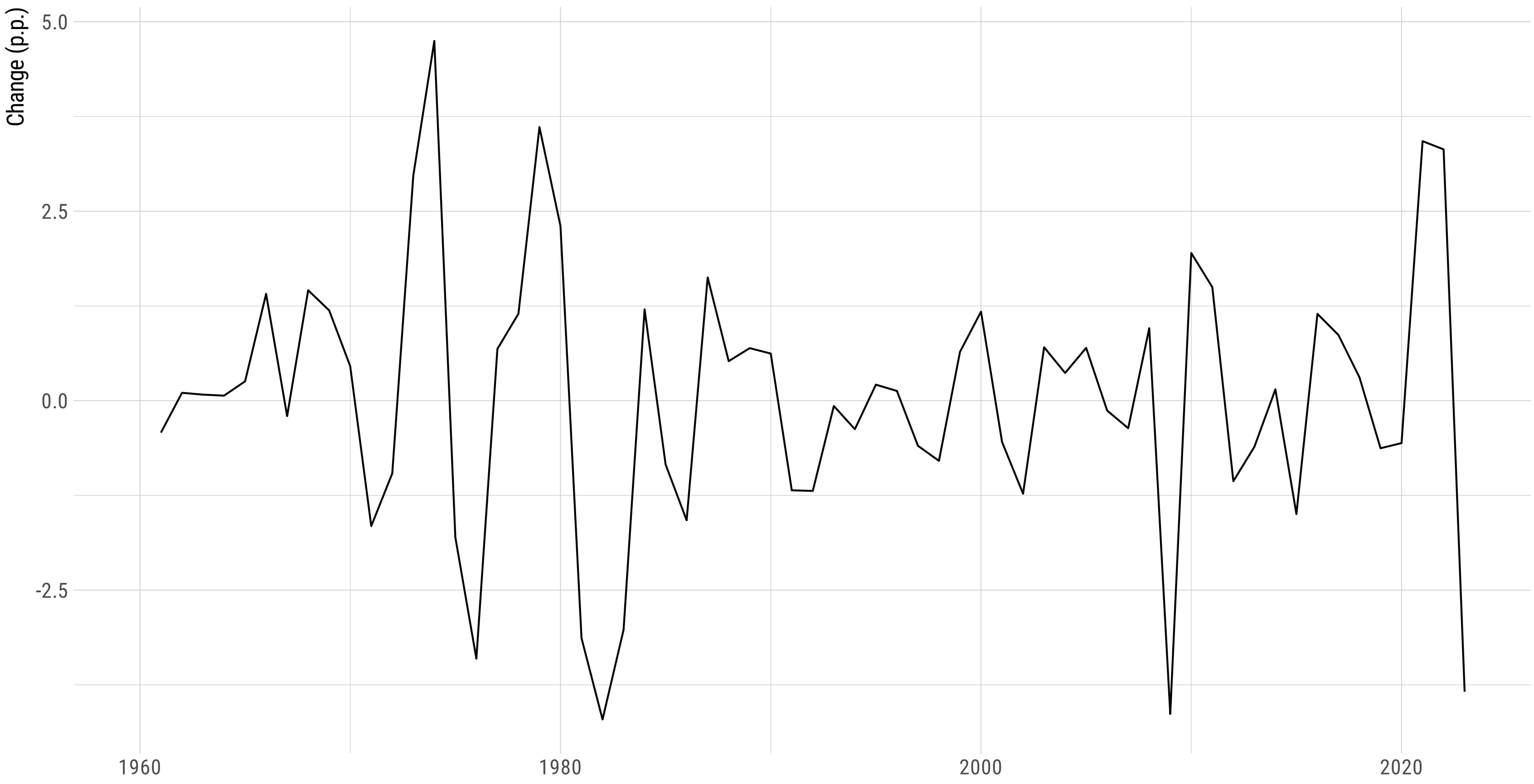


Source: U.S. Bureau of Labor Statistics.

An example: Phillips curve

U.S. annual change in inflation rate (CPI)

1960—2023



Source: U.S. Bureau of Labor Statistics.

An example: Phillips curve

- Are the two variables **stationary**?

```
phillips_ts |>
  features(unrate, unitroot_kpss)
```

```
#> # A tibble: 1 × 2
#>   kpss_stat kpss_pvalue
#>       <dbl>       <dbl>
#> 1     0.109         0.1
```

```
phillips_ts |>
  features(delta_infrate, unitroot_kpss)
```

```
#> # A tibble: 1 × 2
#>   kpss_stat kpss_pvalue
#>       <dbl>       <dbl>
#> 1    0.0763         0.1
```

An example: Phillips curve

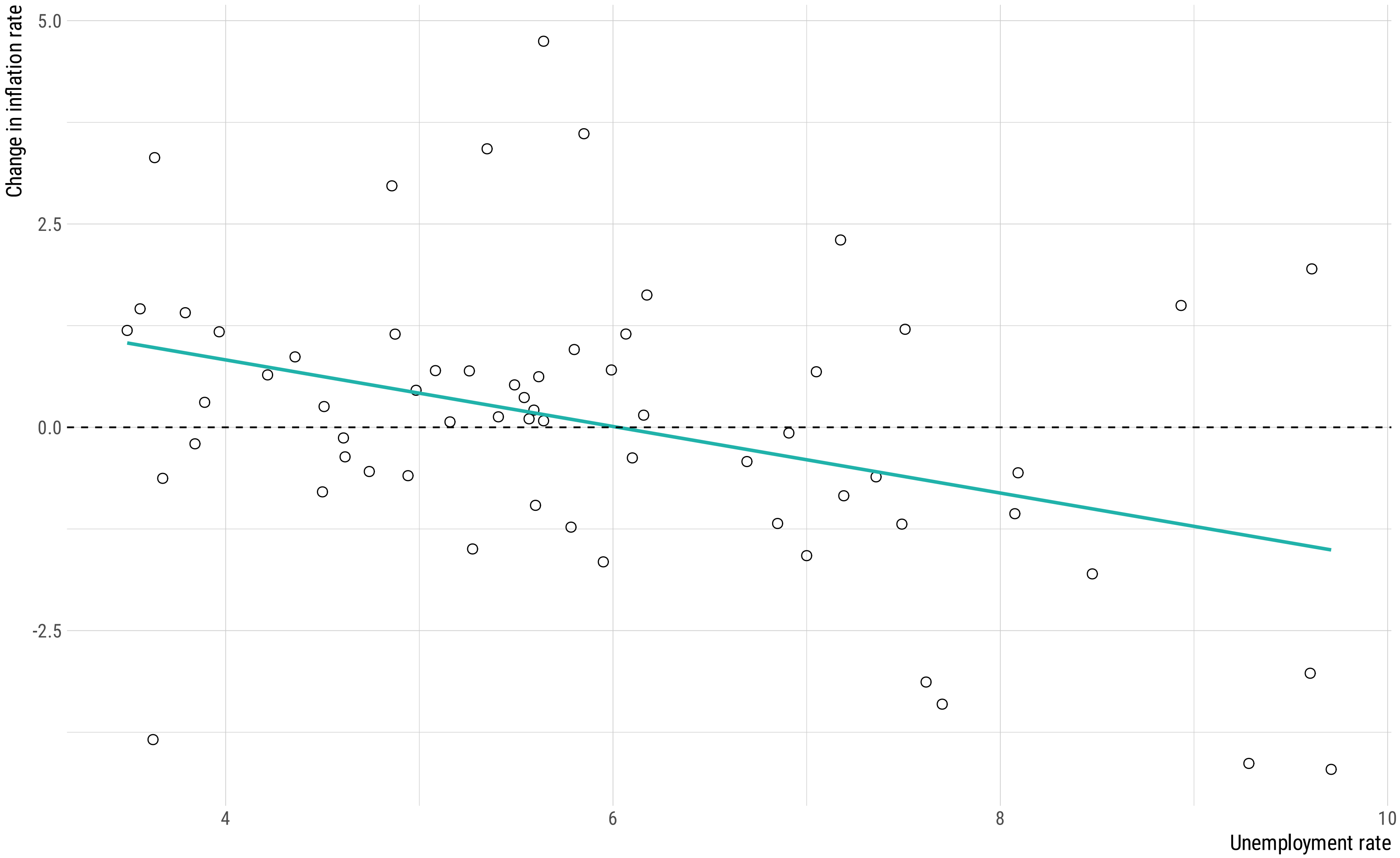
Let us first estimate a **standard** time-series regression model, using the `TSLM()` function.

```
phillips_reg ← phillips_ts ▷
  model(reg = TSLM(delta_infrate ~ unrate))

phillips_reg ▷
  report()
```

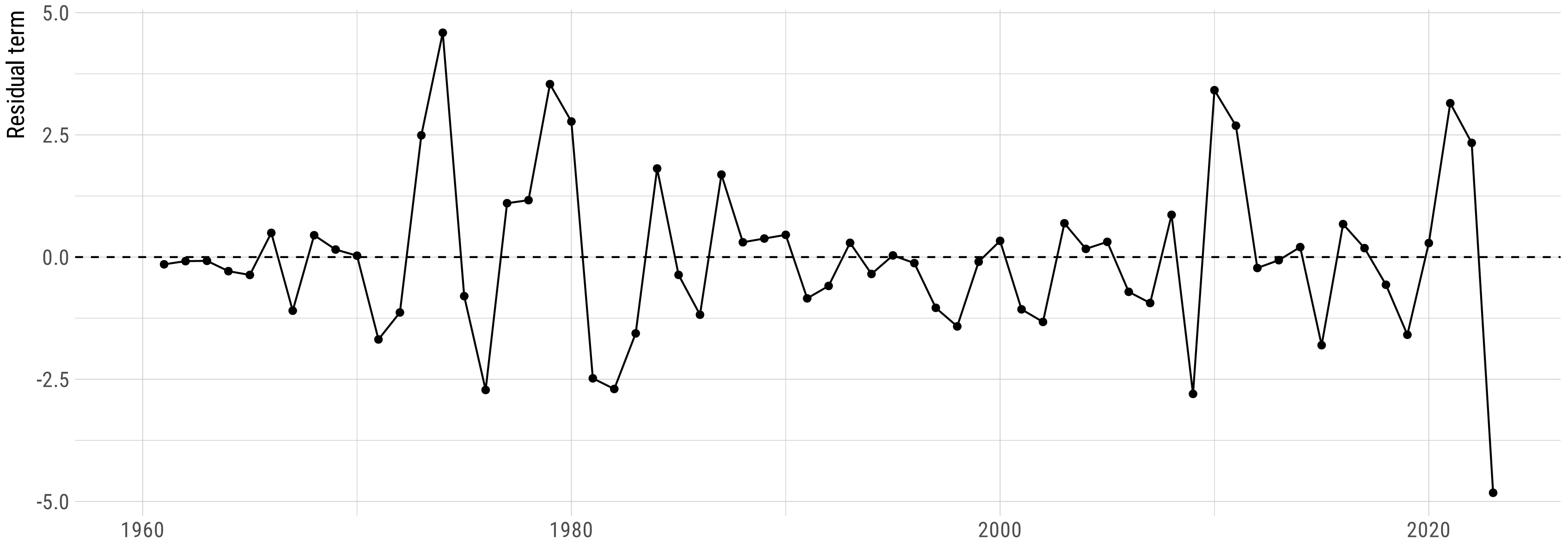
```
#> Series: delta_infrate
#> Model: TSLM
#>
#> Residuals:
#>      Min       1Q   Median       3Q      Max
#> -4.82262 -0.98911 -0.07695  0.47488  4.59029
#>
#> Coefficients:
#>             Estimate Std. Error t value Pr(>|t|)
#> (Intercept)   2.4664     0.7992   3.086  0.00305 **
#> unrate       -0.4093     0.1302  -3.144  0.00257 **
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#>
#> Residual standard error: 1.671 on 61 degrees of freedom
#> Multiple R-squared: 0.1395,	Adjusted R-squared: 0.1254
#> F-statistic: 9.887 on 1 and 61 DF, p-value: 0.0025712
```

An example: Phillips curve

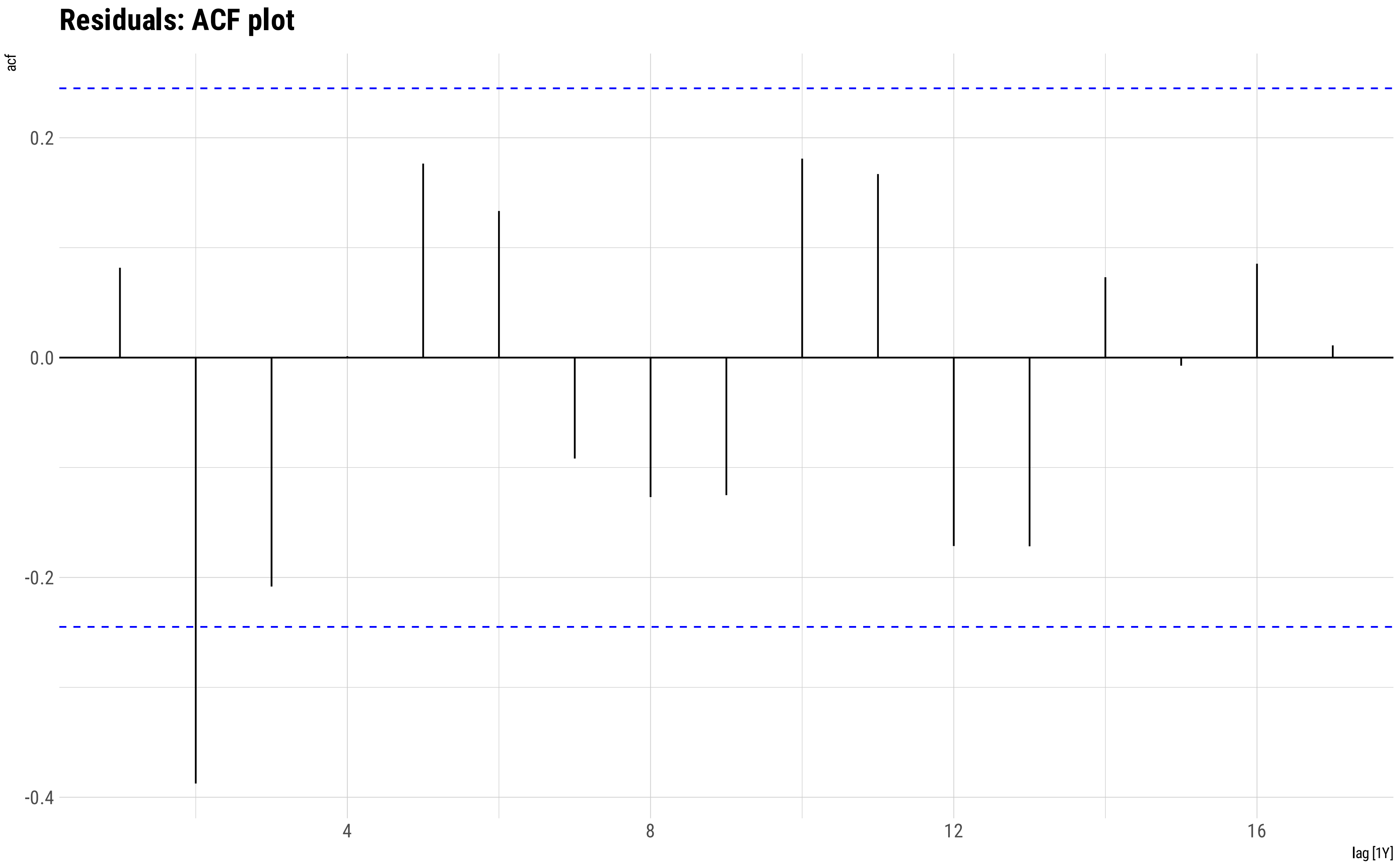


- And now we check the regression's residual term, $\hat{\varepsilon}_t$:

```r
phillips_reg |>
  augment() |>
  select(.innov) |>
  autoplot() +
  geom_hline(yintercept = 0, linetype = 2) +
  geom_point() +
  labs(y = "Residual term",
       x = "") +
  easy_y_axis_title_size(13)
```

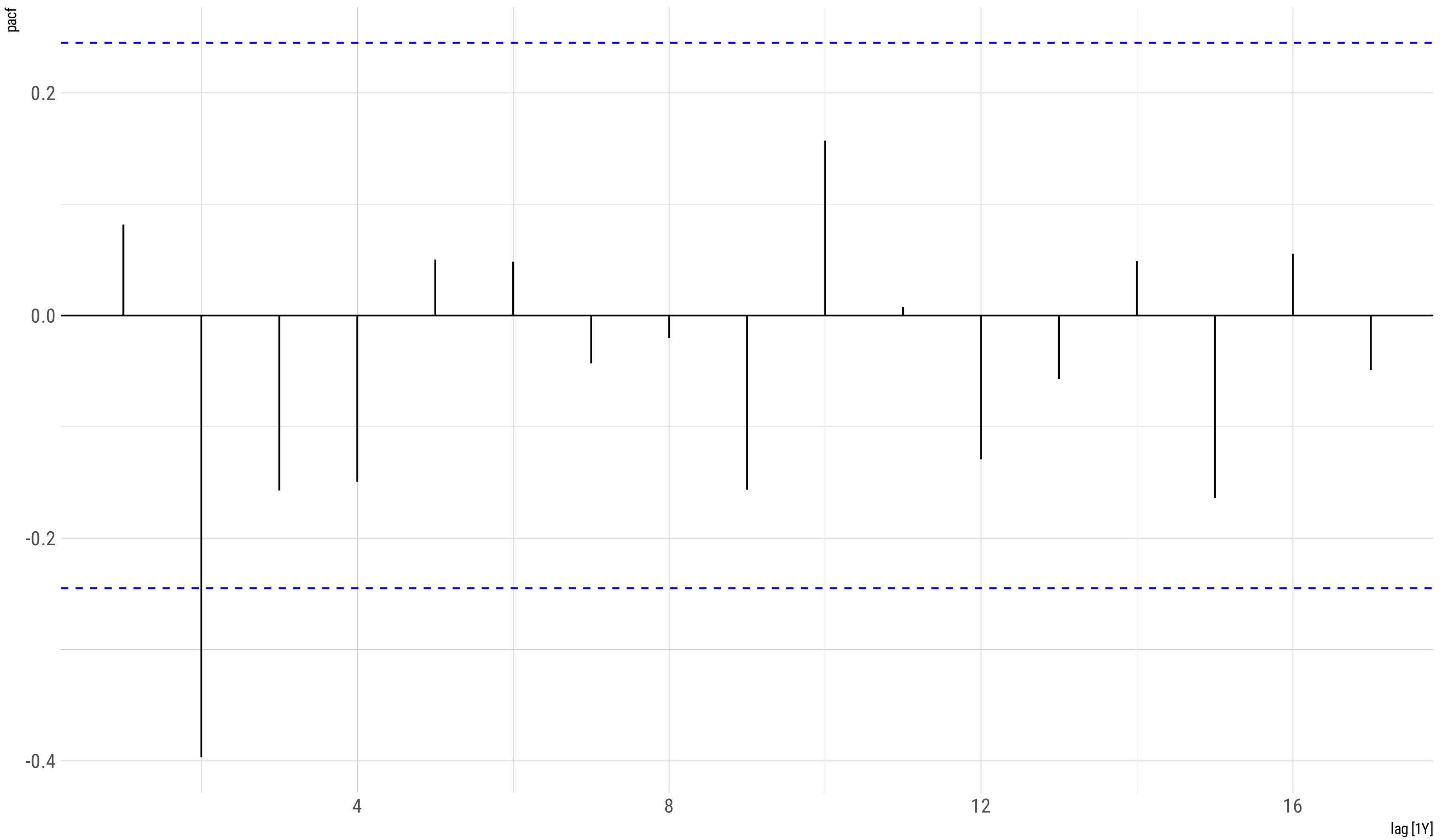


An example: Phillips curve



An example: Phillips curve

Residuals: PACF plot



An example: Phillips curve

- Testing for **residual autocorrelation**:

```
phillips_reg |>
  augment() |>
  features(.innov, ljung_box, lag = 10)
```

```
#> # A tibble: 1 × 3
#>   .model lb_stat lb_pvalue
#>   <chr>    <dbl>     <dbl>
#> 1 reg       22.5    0.0128
```

What do we conclude?

An example: Phillips curve

When the residual term of a regression shows evidence of **serial correlation (autocorrelation)**, inference is **unreliable**.

For **forecasting** purposes, *prediction intervals* will be incorrect.

Therefore, we must **incorporate** the autocorrelation in the residuals into our regression model.

- This is done through an **ARIMA modeling** of this term.

An example: Phillips curve

As the two model variables are **stationary**, *no differencing* is required.

For **ARIMA residuals**, we can use the `ARIMA()` function in a regression context.

```
phillips_arima <- phillips_ts |>
  model(arima_reg = ARIMA(delta_infrate ~ unrate))
```

An example: Phillips curve

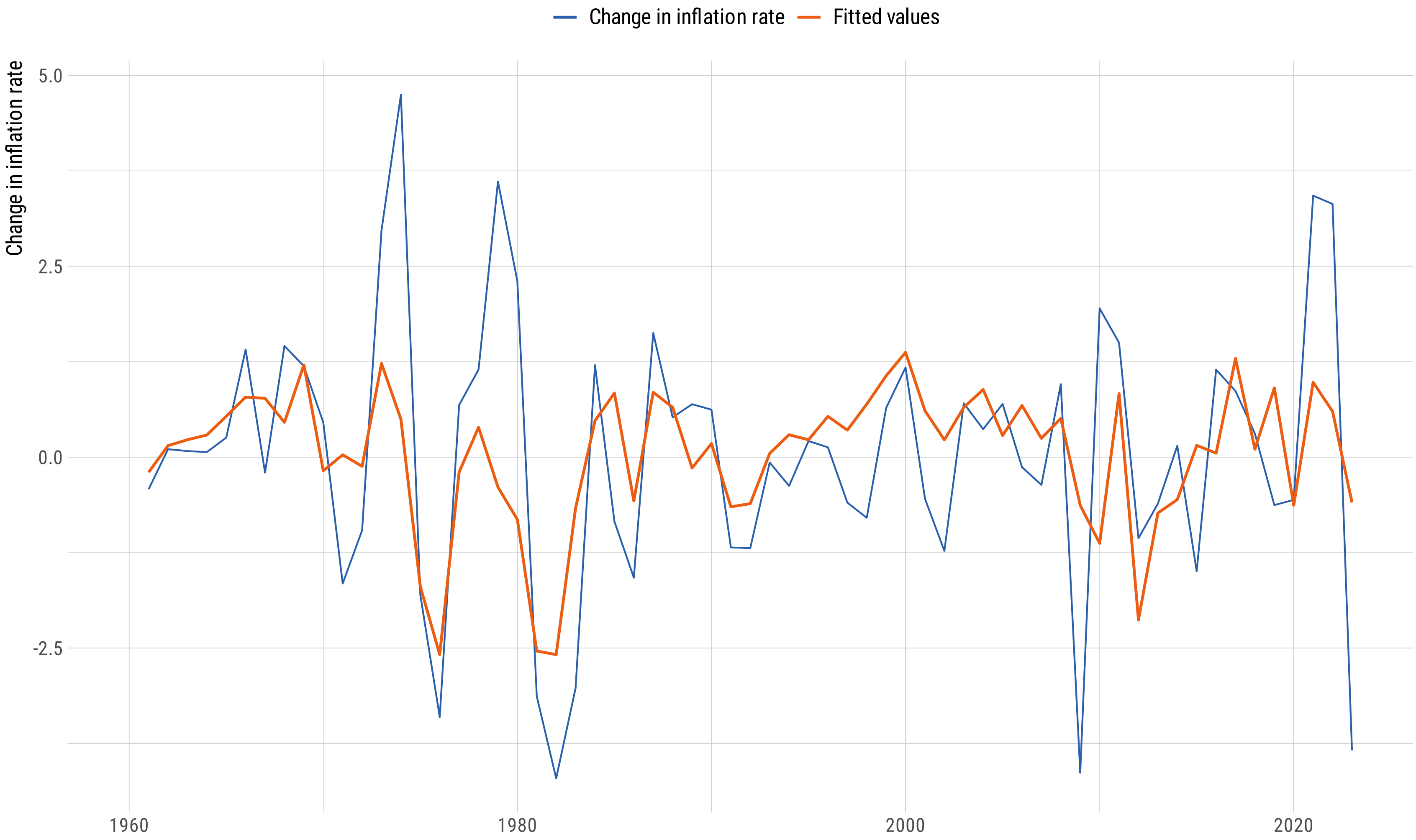
```
phillips_arima ← phillips_ts ▷
  model(arima_reg = ARIMA(delta_infrate ~ unrate))
```

```
phillips_arima ▷
  report()
```

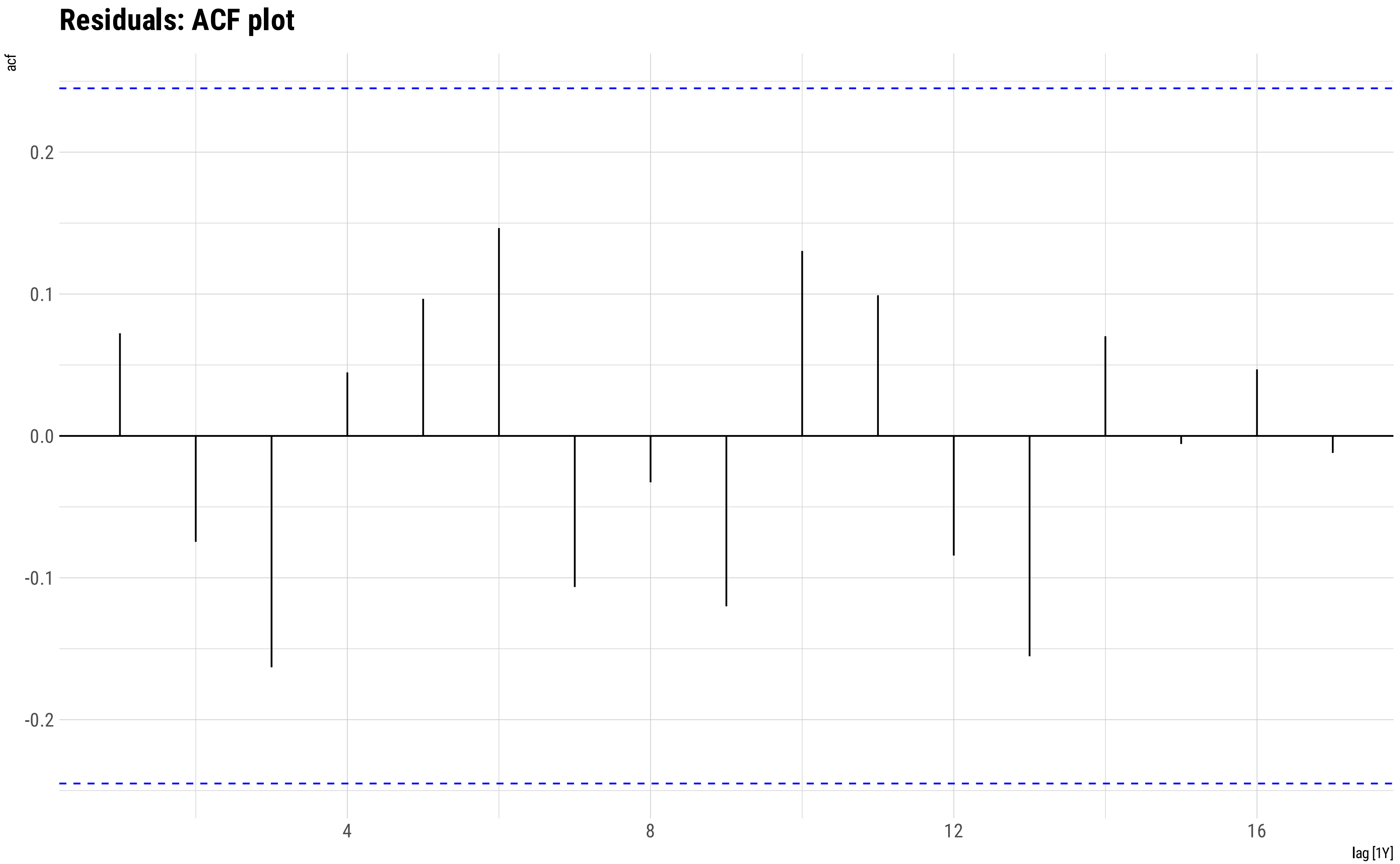
```
#> Series: delta_infrate
#> Model: LM w/ ARIMA(0,0,2) errors
#>
#> Coefficients:
#>           ma1      ma2   unrate  intercept
#>       -0.0317  -0.4999  -0.2891     1.7626
#> s.e.   0.1068   0.1018   0.0822     0.4998
#>
#> sigma^2 estimated as 2.229:  log likelihood=-113.39
#> AIC=236.78   AICc=237.81   BIC=247.57
```

Let us **write out** the estimated model.

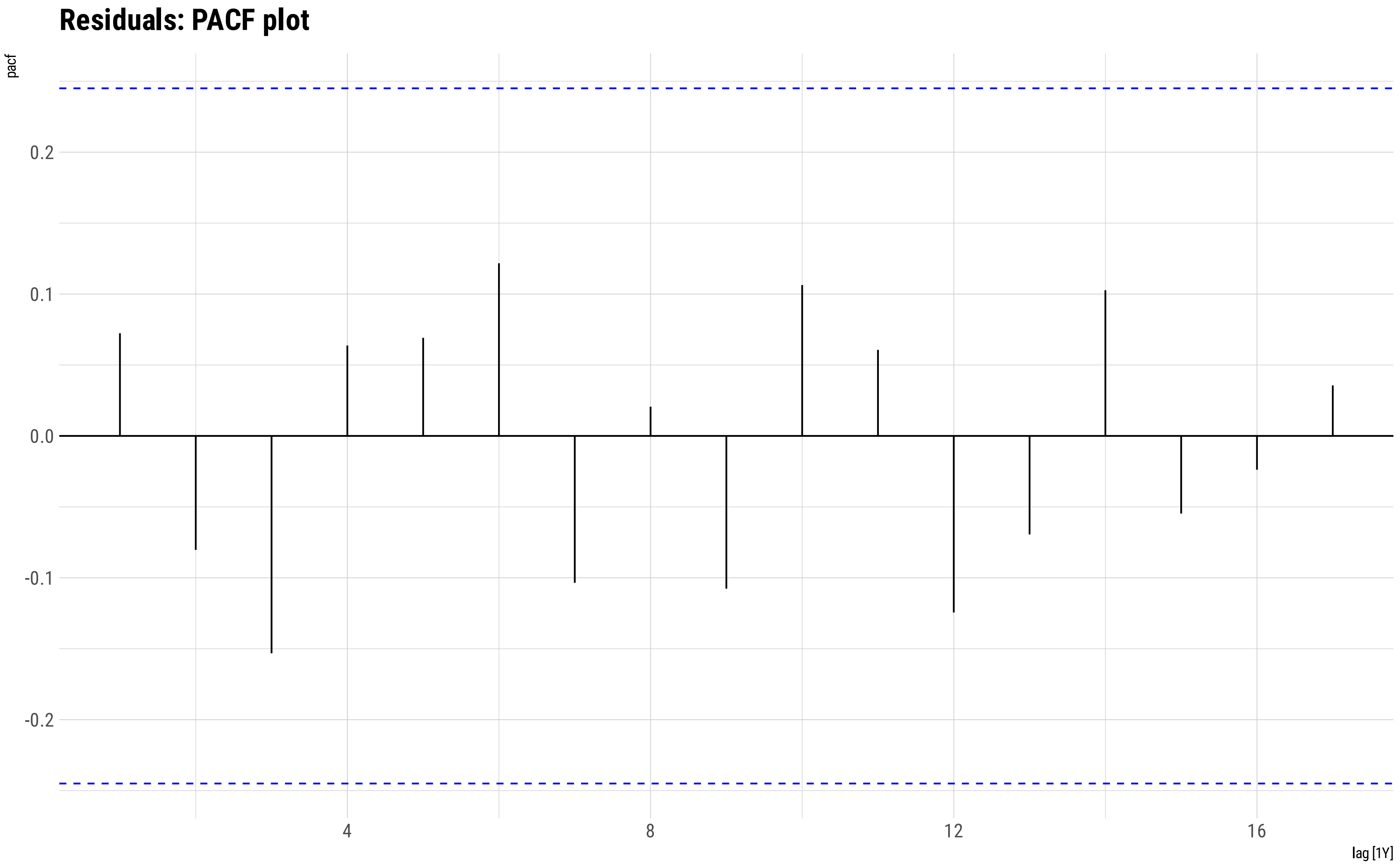
An example: Phillips curve



An example: Phillips curve



An example: Phillips curve



An example: Phillips curve

- Testing for serial correlation:

```
phillips_arima |>
  augment() |>
  features(.innov, ljung_box, lag = 10)
```

```
#> # A tibble: 1 × 3
#>   .model    lb_stat lb_pvalue
#>   <chr>       <dbl>     <dbl>
#> 1 arima_reg    8.19     0.610
```

What do we conclude?

Forecasting

Forecasting

When estimating dynamic regression models with **ARIMA errors**, we need to forecast the regression part of the model and the ARIMA part of the model, and **combine** the results.

When the independent variables are **known** into the future (e.g., *seasonal dummy variables*, *trend* component), their future values are *known*.

However, when this is **not** the case, the usual approach is to **assume future values** for each predictor variable.

In the case of our Phillips curve example, it is **impossible** to know the unemployment rate in the future.

- Thus, we need to make *assumptions*.

Forecasting

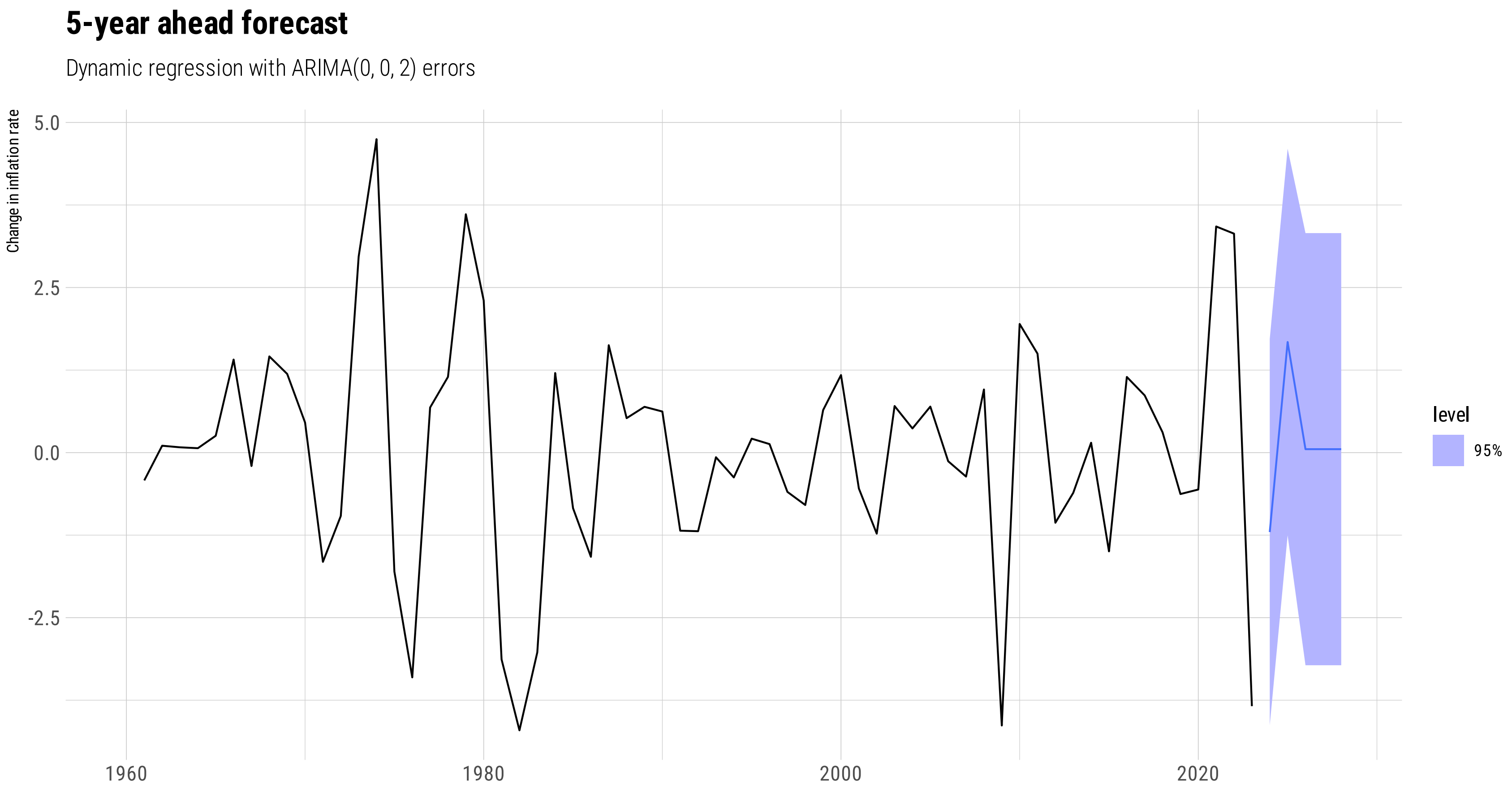
For our example, suppose our forecast horizon is $h = 5$ years.

We may assume that the unemployment rate for this future years will be equal to its **historical mean**.

```
phillips_new_data ← new_data(phillips_ts, n = 5) ▷
  mutate(unrate = mean(phillips_ts$unrate))
```

```
phillips_fc ← phillips_arima ▷
  forecast(h = 5, new_data = phillips_new_data)
```

```
phillips_fc ▷
  autoplot(phillips_ts, level = 95) +
  labs(title = "5-year ahead forecast",
       subtitle = "Dynamic regression with ARIMA(0, 0, 2) errors",
       y = "Change in inflation rate", x = "")
```



Next time: Final thoughts on dynamic regression models