

Granger Causality

Module 9 · What did the added predictor buy us?

Gary Cornwall

Econ 6376 · The George Washington University

Part I • The predictive question

An added regressor can be informative without being a policy lever.

Where we are

$$y_t = \alpha + \sum_{j=1}^p \phi_j y_{t-j} + \sum_{\ell=1}^q \beta_{\ell} x_{t-\ell} + \epsilon_t$$

Mixing-board channel	Today
Intercept and outcome memory	Retain
Observed predictor history	Ask whether it adds information
Current x_t , trend, explicit MA terms	Off in this baseline

Module 8: specify an ARDL. **Module 9:** compare its predictive information.

What did the added predictor buy us?

This guy again!

Clive Granger

A series $x_{1:T}$ is said to “Granger cause” $y_{1:T}$ if past values of x_t have additional power in forecasting y_t , after controlling for the past of y_t .

How would we turn that question into a regression test?

In the ARDL for y with q lags of x :

- $H_0: \beta_1 = \dots = \beta_q = 0$, written $H_0: x \nrightarrow y$
- H_1 : at least one $\beta_\ell \neq 0$, written $H_1: x \rightarrow y$



Incremental, directional, conditional

Word	What it adds to the question
Incremental	Does x help beyond the information already included?
Directional	Predicting y from x differs from predicting x from y .
Conditional	The answer depends on the other information available.

Our implementation: **one-step, linear prediction with specified finite lags.**

Name the outcome, horizon, and information set before naming the result.

A useful signal need not be a lever

An early report can reveal news that also changes the outcome.

- The report helps a forecaster who has not observed that news.
- Changing the recorded report need not change the underlying event.
- Predictive evidence alone does not tell us what an intervention would do.

Would changing an initial-claims report itself change the number of unemployed people?

What happens within the period?

$$y_t = \alpha + \sum_{j=1}^p \phi_j y_{t-j} + \underbrace{\beta_0 x_t}_{\text{same period}} + \underbrace{\sum_{\ell=1}^q \beta_\ell x_{t-\ell}}_{\text{past } x: \text{ the Granger question}} + \epsilon_t$$

- The Granger test keeps only $\ell \geq 1$. Granger called the β_0 channel **instantaneous causality**.
- A past-only test can miss a real within-period effect.
- Same-period dependence alone does not reveal which variable moved first.
- Coarser sampling (monthly rather than weekly) pushes more responses into the $\beta_0 x_t$ term.

If claims respond to a layoff wave within the same month as unemployment, which term picks it up?

Expectations can reverse the apparent story

Investors receive information about future dividends.

The **stock price level today** can incorporate that information before **dividends tomorrow** are observed.

- Past prices help predict dividends, although prices do not affect dividends.
- Past dividends do not help predict prices, although expected dividends set prices.

Which direction would the prediction test label, and what would that label mean?

Part II • The joint test

Keep the regression comparison fixed and remove one lag block.

Restricted and unrestricted fits

Unrestricted

$$y_t = \hat{\alpha}_U + \sum_{j=1}^p \hat{\phi}_{U,j} y_{t-j} + \sum_{\ell=1}^q \hat{\beta}_{U,\ell} x_{t-\ell} + e_{U,t}$$

Restricted

$$y_t = \hat{\alpha}_R + \sum_{j=1}^p \hat{\phi}_{R,j} y_{t-j} + e_{R,t}$$

Only the included x -lag block is removed. Re-estimate the remaining coefficients.

The whole lag block

$$H_0: \beta_1 = \beta_2 = \dots = \beta_q = 0$$

The alternative allows **at least one** included cross-lag coefficient to differ from zero.

If three predictor coefficients sum to zero, have we imposed this null?

No. Offsetting coefficients can still change predictions. The null sets each included coefficient to zero.

The same comparison

Hold fixed

Why?

Outcome rows and transformations

RSS values must refer to the same target observations.

Own lags and other controls

Isolate the added predictor history.

Deterministic terms

Preserve the intended nested models.

Build the complete lagged sample **before** fitting either regression.

What goes wrong if the restricted fit uses four more months?

Granger causality via an F-test

$$F = \frac{(RSS_R - RSS_U) / q}{RSS_U / (n - k_U)}$$

Quantity	Meaning
q	Number of excluded predictor coefficients
n	Common usable observations
k_U	All unrestricted regression coefficients, including the intercept

Reference: $F(q, n - k_U)$, **approximately under suitable dynamic-regression conditions.**

Count first, then calculate

$n = 120$. The unrestricted model has an intercept, two own lags, and three x lags.

$$k_U = 6, \quad q = 3, \quad n - k_U = 114$$

With $RSS_R = 108$ and $RSS_U = 90$,

$$F = \frac{(108 - 90) / 3}{90 / (120 - 6)} = 7.6.$$

```
1 pf(7.6, df1 = 3, df2 = 114, lower.tail = FALSE)
```

```
[1] 0.0001118079
```

An ARDL in action

For two own lags and three predictor lags:

```
1 restricted <- lm(y ~ y_l1 + y_l2,  
2               data = common, na.action = na.fail)  
3 unrestricted <- lm(y ~ y_l1 + y_l2 + x_l1 + x_l2 + x_l3,  
4               data = common, na.action = na.fail)  
5 anova(restricted, unrestricted, test = "F")
```

`common` contains the correctly dated lag columns and the shared complete sample.

Read the fitted models before reading the p-value.

Reverse the direction with another equation

$$x_t = a + \sum_{j=1}^{p_x} \rho_j x_{t-j} + \sum_{\ell=1}^{q_y} \gamma_{\ell} y_{t-\ell} + \epsilon_{x,t}$$

Test $H_0: \gamma_1 = \dots = \gamma_{q_y} = 0$ in this separate scalar regression.

- Evidence can appear in one direction or both.
- A nonrejection can reflect weak information, a small sample, or an unsuitable model.
- Neither directional pattern identifies policy effects by itself.

An in-sample test of a predictive idea

The F-test judges a **predictive** idea by **in-sample** fit. [Why not test it out of sample?](#)

- Ashley, Granger & Schmalensee (1980): rely primarily on out-of-sample forecasts.
- Trade-off: in-sample fit can reward overfitting; a pre-chosen holdout costs power.

Ashley & Tsang (2014)

Split at τ ; fit each side, predict the other; repeat over τ

Pseudo-F from prediction errors, with bootstrap inference

Cornwall et al. (2019)

Hold out a block; fit on the rest, minus the next p rows

Posterior simulation of RMSPE/MAE; posterior odds

Both: every prediction uses data not used to fit that model, and nearly all the data are used.

Part III • Generate the evidence

We know the data-generating mechanism. The test sees one sample.

Generating some data

$$x_t = 0.55x_{t-1} + \epsilon_{x,t}$$

$$y_t = 0.4y_{t-1} + 0.6x_{t-1} + \epsilon_t$$

```
1 for (t in 2:total) {  
2   x[t] <- 0.55 * x[t - 1] + epsilon_x[t]  
3   y[t] <- 0.4 * y[t - 1] + 0.6 * x[t - 1] + epsilon[t]  
4 }
```

Independent standard-normal innovations. Retain 600 observations after burn-in.

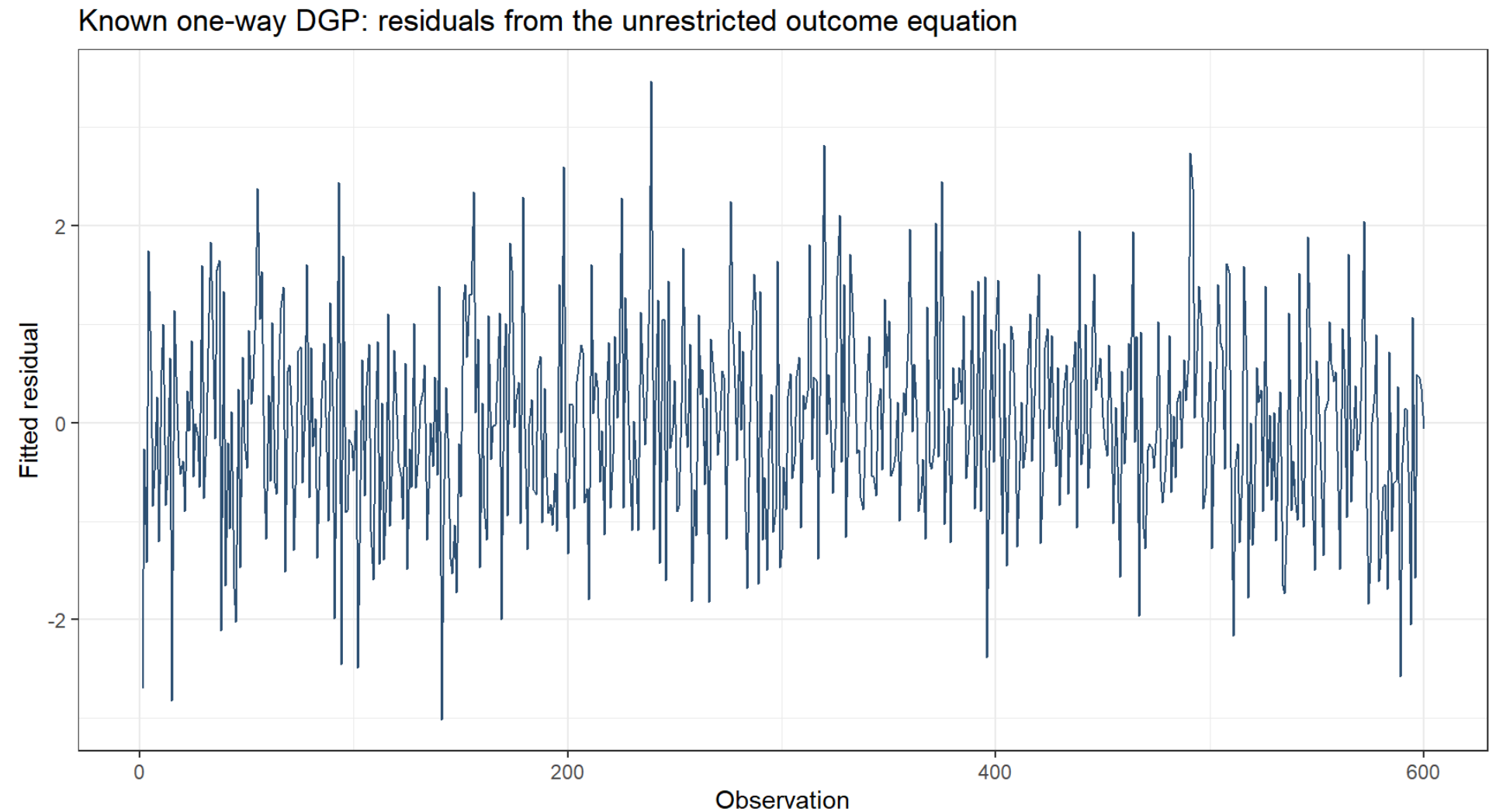
Which history should help predict which outcome?

Comparing the mechanism and the sample

Added history	F	p-value
x in the y equation	259.355	$< 10^{-6}$
y in the x equation	1.064	0.303

The true cross-lag coefficient is **0.6**. Its estimate rounds to **0.600**.

Change 0.6 to zero, keep the innovations, and re-estimate. Then shorten the sample.



Fitted residuals over the 600 observations from the known one-way simulation.

An early signal of the same news

Let z_t be news and x_t a noisy report of it:

$$x_t = z_t + \eta_t, \quad y_t = 0.4y_{t-1} + 0.8z_{t-1} + \epsilon_t.$$

The three disturbances are mutually independent standard-normal draws over time.

There is **no x term** in the structural equation for y .

Can yesterday's report still improve today's prediction?

What changes when we observe the news?

Information already available	Add x_{t-1} ?
Outcome history only	The report reveals otherwise hidden news.
Outcome history and z_{t-1}	The news itself is already known.

What happens to the population predictive contribution of x ?

It disappears in the second comparison. Its conditional population coefficient is zero.
The two regressions answer different questions.

A true null can still be rejected

Information already available	Sample p-value	Population restriction
Own history only	$< 10^{-6}$	x adds predictive information.
Own history plus news	0.0157	x adds no further information.

The second test rejects at 5%. **It is a false rejection in this simulated sample.**

The first finding is genuine predictive information, even though changing the recorded signal does not change the outcome mechanism.

Part IV • Claims and unemployment

Use the Module 8 data, then report the uncertainty the model leaves behind.

What exactly are we predicting?

$$y_t = \Delta U_t, \quad x_t = 100\Delta \log C_t$$

Object	Definition
U_t	Monthly unemployment rate, in percent
C_t	Mean of weekly initial claims whose week-ending Saturdays fall in month t
y_t	Change in unemployment, in percentage points
x_t	Approximately the percentage change in that monthly claims mean

Why would claims contain information beyond unemployment's own history?

Dates on the data and dates on the clock

- **Training:** January 1990–December 2017, 336 usable observations.
- **Evaluation:** January 2018–December 2019, 24 observations.
- Prediction begins after both previous-month measures have been released, before the target month's unemployment release.

This is usually a **nowcast of the current calendar month**.

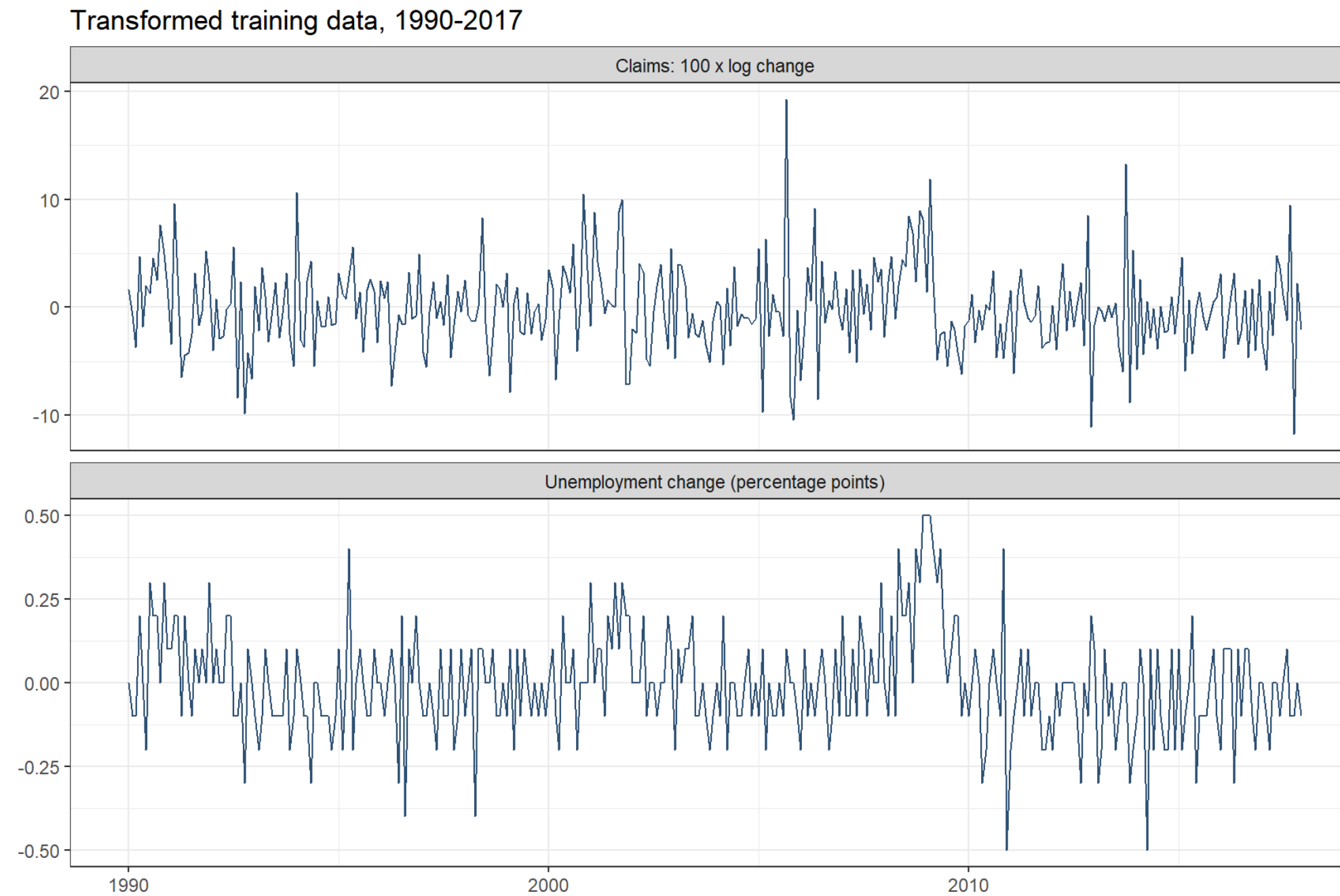
Revised historical caches support a pseudo-out-of-sample exercise, not a real-time vintage evaluation.

Timing can create or hide predictive direction

Data feature	In this exercise	Why it matters
Frequency	C_t averages weekly claims before and after the mid-month CPS reference week	A monthly “lag” is not a clean calendar lead
Seasonal adjustment	BLS and DOL adjust each series separately	Filtering can add or remove lagged correlation
Release timing	Claims for month t 's reference week are public about two weeks before U_t	Our past-only design leaves them out on purpose

The forecast comparison therefore **understates** claims' practical nowcasting value. Using x_t is the same-period question from Part I.

A quick look at the data



Training data from 1990 through 2017: changes in unemployment and log changes in the monthly mean of weekly claims, on separate scales.

Working in changes is a specification choice. The stationarity checks and their qualifications are in the notes.

Fix the comparison before reading the result

Six own lags and three predictor lags in each outcome equation.

For unemployment change:

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

Count	Value
Common observations n	336
Restrictions q	3
Unrestricted coefficients k_U	10
Residual degrees of freedom	326

The main specification is inherited from Module 8's exploratory work; it is not a fresh prespecified test.

The result and the diagnostic belong together

Added history and outcome	F	Nominal p	BG(12) p
Claims growth in unemployment change	21.058	1.71×10^{-12}	0.0031
Unemployment change in claims growth	1.804	0.1463	0.2549

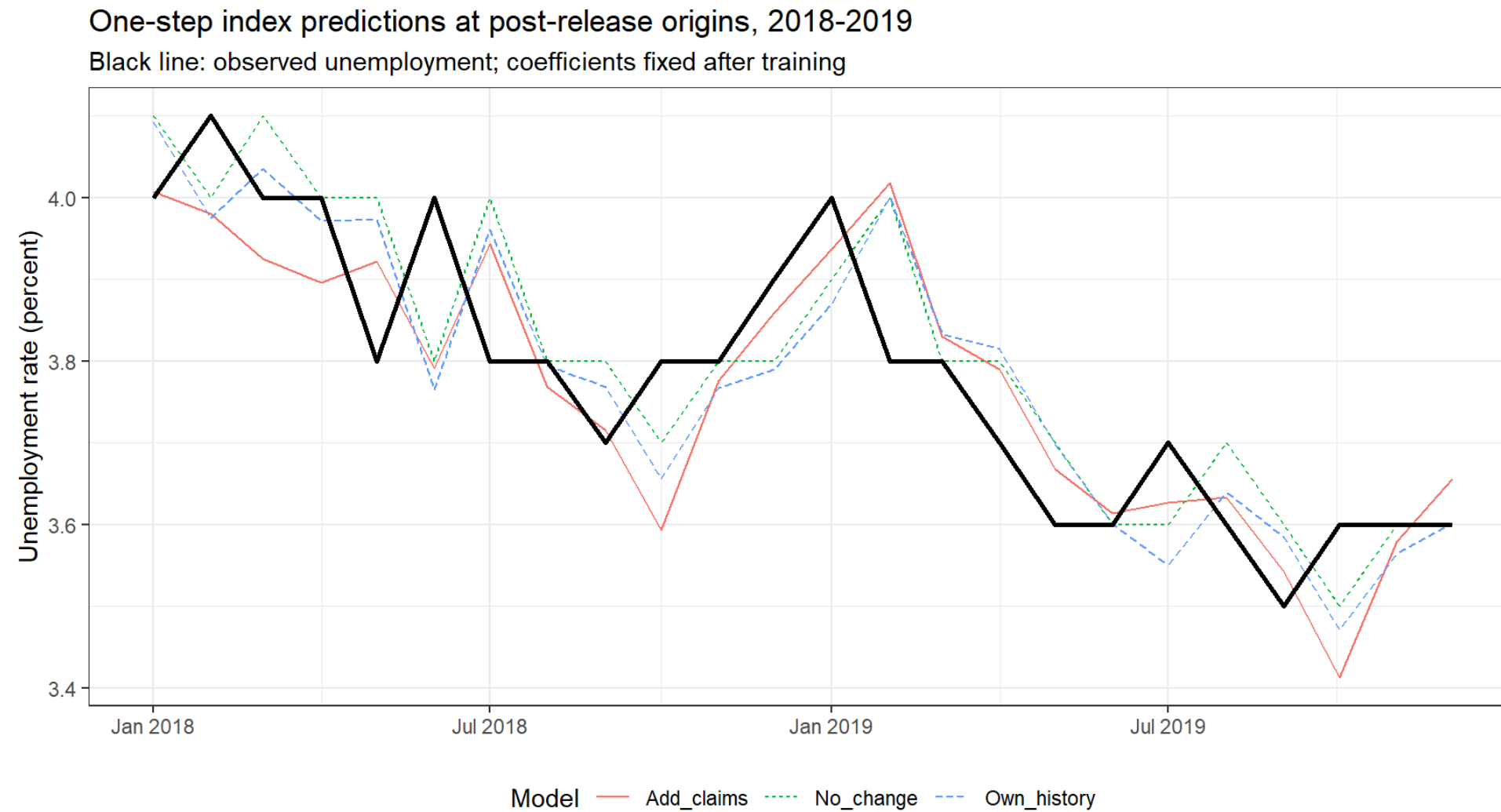
BG(12) checks residual serial correlation through twelve lags.

The reverse result depends on lag choice. Nominal p-values with $q = 1$: 0.00518 at $(p, q) = (3, 1)$ and 0.0286 at $(6, 1)$; at the main $(6, 3)$, 0.146. **Fix lag lengths in advance.**

How much weight can the tiny forward p-value carry when residual dynamics remain?

The forward specification fails this diagnostic. Report the nominal calculation and limitation; a **validated Granger-causality verdict is not established here.**

Forecast usefulness is a separate comparison



Observed unemployment and three one-step prediction series for 2018 and 2019, with fixed training coefficients.

2018–2019 RMSE

Model	RMSE
Own history	0.1129
Add claims	0.1057
No change	0.1099

Unemployment percentage points.

A modest improvement in this 24-month evaluation.

A lower in-sample RSS is mechanical. A lower held-out loss is an empirical result.

What would you report?

1. The transformed outcome, units, dates, horizon, and available information.
2. The complete lag block excluded, common sample, F statistic, and nominal p-value.
3. The specification choices and diagnostics that support or limit interpretation.
4. The separate held-out comparison, if one was made.

Rewrite “Claims cause unemployment, so reducing claims would reduce unemployment.”

Part V • Uses and limits

The predictive question travels across applications. Its interpretation still depends on the setting.

Economics and beyond

Setting	Predictive question	Interpretation issue
Money and income	Does monetary history add information about activity?	Same variables, different answer: yes for 1960s–70s data, no for 1970s–80s.
Unemployment and prices	Does labor-market history improve price prediction?	A predictive result alone does not identify a policy trade-off.
fMRI signals	Does one region's measured history help predict another's?	Blood-oxygen delays and sampling can affect the measured direction.

Granger causality has uses beyond economics. It retains the same information-based question.

A research connection

Cornwall, Mills, Sauley, and Weng (2019)

Predictive Testing for Granger Causality via Posterior Simulation and Cross-Validation

The Phillips-curve application finds some in-sample rejections alongside weak held-out predictive evidence.

What was predicted, against which benchmark, on which observations, and under which loss?

The method is on the Part II slide “An in-sample test of a predictive idea.” Coefficient restrictions and held-out performance measure different evidence.

Key takeaways

Granger causality concerns added predictive information relative to a stated history.

Test the whole included predictor-lag block, on the same observations, using a suitable model.

A test decision can be wrong. A nonrejection is not proof of no relationship.

Diagnostics and held-out prediction matter, and neither identifies an intervention effect by itself.

Call these tests of whether x helps forecast y , not whether x causes y . — after Hamilton (1994)

Next time: cointegration and error correction

Routine lag-exclusion inference can fail for integrated levels.

Differencing everything can also discard information about a shared long-run relationship.

Module 10: cointegration and error correction, through the same scalar ARDL.

What if two series wander, but their separation stays stable?

Preview only. The notes' optional "Looking Ahead" section supplies the algebra.