

Vector autoregression with real data: inflation, unemployment and the Fed's rate

Fitted to 1960-2019, the classic three-variable VAR put US inflation in the second quarter of 2022 at 3.2%, with an 80% band reaching 6.9%. It came in at 9.3%. This guide fits that model to real data, reads it, tests it, and shows what such a model can and cannot tell you, from monetary-policy shocks to the odds for 2027.

Functions	<code>fx.RiskVARSelectλ</code> · <code>VARFitλ</code> · <code>VARSummaryλ</code> · <code>VARResidsλ</code> · <code>VARStabilityλ</code> · <code>VARPortmanteauλ</code> · <code>Grangerλ</code> · <code>VARIRFλ</code> · <code>VARIRFBandsλ</code> · <code>VARFEVDλ</code> · <code>VARForecastλ</code> · <code>VARSimλ</code> · <code>VARBacktestλ</code> , with <code>Exog</code> and <code>FutureExog</code>
In XL Edge	XL Edge tab > Productivity Tools > Stat Tools > Insert Time Series > Multivariate - VAR and cointegration
Example	<i>XL Edge Example - VAR 2 Real Macro Data.xlsx</i> (same folder). Every number in this guide is read from it.
Data	FRED, Federal Reserve Bank of St. Louis: CPIAUCSL and UNRATE (U.S. Bureau of Labor Statistics), FEDFUNDS (Federal Reserve Board). Public domain; downloaded 2026-10-18 17:14 UTC, January 1959 to September 2026.
Series	Part 2 of 2. Part 1, <i>VAR 1 Business Forecasting</i> , introduces every function on simulated data whose true model is known; this part assumes it.

- | | |
|---|--|
| 1 The classic VAR, and what is different here | 8 Step 5. Who leads whom? |
| 2 The data, holes included | 9 Step 6. A monetary-policy shock |
| 3 Step 0. Levels or differences? | 10 Step 7. The 2019 forecast against what happened |
| 4 Step 1. Which sample? | 11 Step 8. Backtests against simple benchmarks |
| 5 Step 2. How many lags? | 12 Step 9. Model B: all the data, and 2026-2027 |
| 6 Step 3. Fit Model A and read it | 13 Lessons from real data |
| 7 Step 4. Diagnostics that fail | 14 Checking the model, and function notes |

1 The classic VAR, and what is different here

In 1980 Christopher Sims proposed letting a few macroeconomic series explain one another, without a theory saying in advance which causes which. Twenty years later James Stock and Mark Watson used three series to show what such a model can do: **inflation**, the **unemployment rate** and the **federal funds rate**, the interest rate the Federal Reserve sets. Those three still frame most discussions of monetary policy. This guide fits the same three to US data from 1960 to 2025.

Part 1 worked on simulated data, where the true model was known and every estimate could be graded. Real data offer no such answer key. So the questions change:

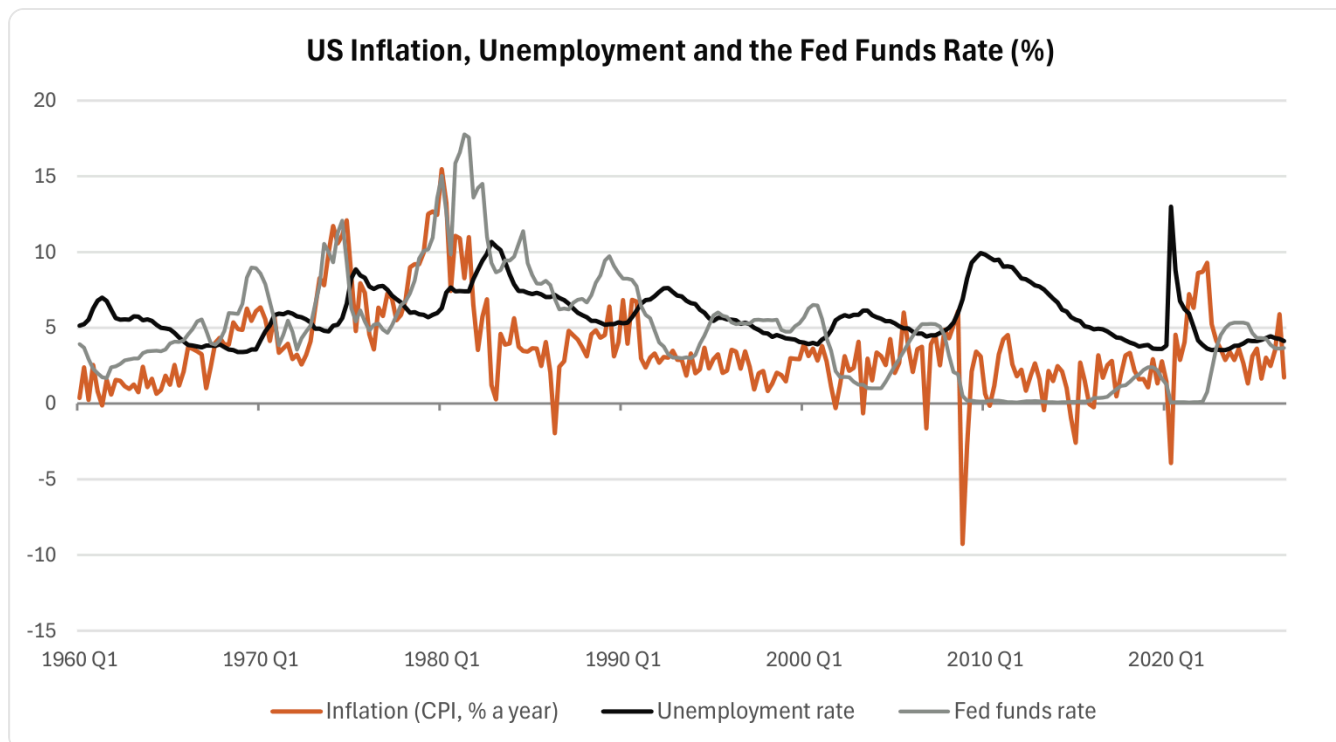
- **Which data?** Real series have holes, revisions and once-in-a-century quarters.
- **Which model?** The criteria, the diagnostics and the backtests can disagree, and judgment has to fill the gap.
- **How much to believe it?** Without a truth to compare with, the only tests are diagnostics, robustness (does the answer survive another order, sample or lag length?) and forecasts scored against what then happened.

The workbook fits two models. **Model A** uses 1960 to 2019, the classic sample, and its forecast for 2020 to 2025 can be scored. **Model B** uses everything through 2025, with the two COVID quarters handled by *exogenous dummies*, and forecasts 2026 and 2027.

2 The data, holes included

Sheet *FRED monthly* holds the three series as downloaded, from January 1959 to September 2026. Sheet *Data* averages each quarter's three months with ordinary formulas, so every quarterly number can be traced back to its source:

Series	Definition	Name
Inflation	$400 \times \text{LN}(\text{CPI} \div \text{last quarter's CPI})$: the quarter's price change at an annual rate (CPIAUCSL, seasonally adjusted)	infl
Unemployment rate	The quarter's average, % (UNRATE, seasonally adjusted)	unemp
Fed funds rate	The quarter's average, % (FEDFUNDS)	ff



Sixty-six years in one chart (sheet *Data*): the inflation of the 1970s, the rate shocks of 1980-81, the long decline in rates, the zero-rate years after 2008, the COVID spike and the 2021-23 inflation.

REAL DATA HAVE HOLES

No CPI or unemployment rate was published for October 2025: the federal government shutdown stopped the surveys. Something has to go in that cell before the fourth-quarter average can be taken. The workbook uses the geometric mean of September and November for the CPI (324.654) and the average of the two for unemployment (4.45%), in yellow cells you can change. The choice is not innocent: leave October out and average November and December only, and 2025 Q4 inflation is 2.85% rather than 2.48%. That quarter is where Model B's forecast starts. Write such choices down, make them visible, and check what they change.

Stock and Watson measured inflation with the GDP deflator and used data to 2000. This guide uses the CPI and data to 2025, so its numbers are not a replication of theirs, only the same design. FRED also revises the CPI every February, so a later download changes the numbers a little:

`docs\build\fetch_fred.py` refreshes them.

3 Step 0. Levels or differences? fx.RiskUnitRootA

Part 1 differenced its series, because dollar levels that trend have no average to return to. Here the question is harder. Inflation, unemployment and interest rates have no trend, but they are very *persistent*: they wander for years before coming back. On 1960-2019:

Series (in levels)	ADF p	ADF verdict	KPSS p	KPSS verdict
Inflation	0.020	Stationary	0.017	Not stationary
Unemployment	0.045	Stationary	>0.10	Stationary
Fed funds rate	0.330	Unit root	<0.01	Not stationary

Only for unemployment do the two agree on stationarity, and even there Phillips-Perron does not (p 0.12). For the fed funds rate both point to a unit root. For inflation they contradict each other: ADF rejects a unit root, KPSS rejects stationarity. That is typical of sixty years of persistent macro data, which cannot tell "wanders for a decade" from "wanders for ever". Stock and Watson kept all three series in levels, and so does this guide. A VAR in levels stays valid for impulse responses and forecasts even if a series is close to a unit root (Sims, Stock and Watson, 1990), at the price of persistence near 1 (section 7) and less reliable Granger tests (section 8). The alternatives are differencing, which throws away the levels' information, or a VECM if the series are cointegrated (see the *Multivariate Time Series* guide).

4 Step 1. Which sample? COVID and exogenous dummies

In the second quarter of 2020 unemployment averaged 13.0%, up from 3.6% six months earlier, and it fell almost as fast. Nothing in the previous sixty years looks like it. A least-squares fit treats it as the most informative quarter in the sample, because its residual is enormous, and bends the whole model to explain it. Sheet *1 Sample and COVID* fits the unemployment equation three ways:

The unemployment equation	Model A: 1960-2019	1960-2025, no dummies	Model B: 1960-2025 + dummies
Coefficient on last quarter's unemployment	1.585	0.803	1.498
Sum of its own lags (persistence)	0.966	0.893	0.969
Residual standard deviation (points)	0.232	0.688	0.236
Largest root	0.955	0.956	0.962

Two quarters out of 264 almost halve the coefficient on last quarter's unemployment and triple the typical surprise. Every forecast band built on that fit would be far too wide in ordinary times. There are three honest choices: stop the sample in 2019 (Model A); keep everything and let COVID distort the fit; or keep everything and give each extreme quarter its own **dummy**, an exogenous column that is 1 in that quarter and 0 otherwise (Model B):

```
Data!L14:M277, named Covid: one column for 2020 Q2, one for 2020 Q3
=fx.RiskVARFitA(Macro, LagA, , , Covid, HSTACK(MacroNames, {"covid_q2","covid_q3"}))
```

A *pulse dummy* gives its quarter a coefficient of its own, so that quarter is fitted exactly (its residual is zero, a check on sheet *Checks*) and no longer pulls on the other coefficients. With the dummies, the unemployment coefficient is back to 1.498, close to the 1960-2019 value. The dummies' coefficients in the unemployment equation, 9.05 for 2020 Q2 and -8.79 for 2020 Q3, are the size of the shock the model sets aside. Use dummies sparingly and name them in the write-up: each one removes a quarter from the evidence.

5 Step 2. How many lags? `fx.RiskVARSelectA`

```
=fx.RiskVARSelectA(MacroA, 8)
=fx.RiskVARSelectA(Macro, 8, , , Covid) Exog: the dummies sit in every fit
```

Lags chosen by	AIC	HQ	SC (BIC)	FPE
Model A, 1960-2019	3	3	3	3
All data, with the dummies	4	3	2	4

On 1960-2019 all four criteria choose **3 lags**, three quarters of history for every series. On the longer sample they disagree: AIC and FPE want more, SC fewer, HQ sits between, which is the usual pattern. The guide uses 3 lags for both models (the name `LagA`), so they can be compared like with like; section 7's diagnostics are the reason to try more.

6 Step 3. Fit Model A and read it: `fx.RiskVARFitA`, `fx.RiskVARSummaryA`

```
=fx.RiskVARFitA(MacroA, LagA, , , , MacroNames) '3 Fit'!C10, named ModelA
=fx.RiskVARSummaryA(MacroA, ModelA)
```

Model A fits 237 quarters with 3 lags: 28 rows of Model block. The first thing the summary shows is spectacular fit:

	Inflation	Unemployment	Fed funds
R ² in levels (the summary)	0.68	0.98	0.95
R ² of the quarterly change, same residuals	0.33	0.51	0.23
Residual standard deviation (points)	1.74	0.23	0.79

A HIGH R² IN LEVELS IS NOT SKILL

Unemployment's R² of 0.98 sounds like a model that knows the future. It mostly says that unemployment this quarter is close to last quarter's, which anyone could forecast. Measured against the quarterly *change*, the thing a forecast has to get right, the same model, with the same residuals, explains 0.51. For inflation, 0.33. Always ask what an R² is measured against.

Reading the coefficients. The fed funds equation looks like a central bank's reaction: the rate's own lags sum to 0.94 (a slow-moving policy rate), inflation's lags to 0.10 (higher inflation, higher rates) and unemployment's to -0.06 (higher unemployment, lower rates). Because the rate's own lags sum to nearly 1, the long-run effect is the short-run sum divided by 1 - 0.94: the rate settles about 1.6 points higher for each point of inflation and 1.0 lower for each point of unemployment. More than one for one on inflation is the "Taylor principle", found here without being told. Individual coefficients matter less than these sums, because the lags of persistent series are strongly correlated with each other and trade off in the estimates. The residual correlations are moderate (inflation and unemployment -0.07, inflation and the rate 0.27, unemployment and the rate -0.38), so the Cholesky order matters less here than in Part 1, but it still matters (section 9).

7 Step 4. Diagnostics that fail

Test	Result	Reading
Stability: largest root (<code>fx.RiskVARStabilityλ</code>)	0.955	Stable, but a shock's echo halves only in 15.2 quarters
Portmanteau, VAR(3) at 10 lags (<code>fx.RiskVARPortmanteauλ</code>)	p 8.2e-05	Serial correlation
... VAR(8) at 12 lags	p 0.001	Serial correlation: more lags do not cure it
Normality of inflation's residuals: excess kurtosis (<code>fx.RiskNormalityλ</code>)	15.01	Jarque-Bera p < 0.001: very fat tails
Volatility clustering: ARCH(4) on unemployment's residuals (<code>fx.RiskARCHTestλ</code>)	p 1.2e-07	ARCH effects
... on the fed funds rate's residuals	p 7.2e-09	ARCH effects
... on inflation's residuals	p 0.73	No ARCH effects

Part 1's simulated model passed every test because it was the truth. A real one fails most of them, and that is normal, not a reason to give up:

- **Leftover correlation.** Sixty years of policy regimes, oil shocks and financial crises do not follow one set of coefficients. A VAR averages them; the portmanteau test notices.
- **Fat tails.** A few quarters (the oil shocks, 2008) have surprises far beyond a normal curve. The P10-P90 bands, which assume normal shocks, will be about right most of the time and badly wrong some of the time. Pass `DegFree` to `fx.RiskVARSimλ` for Student t shocks if the tails matter to the decision.
- **Changing volatility.** Unemployment and interest rates were far more volatile before the mid-1980s than after (the "Great Moderation"). One Σ for the whole sample makes recent bands too wide and old ones too narrow.

What to do: use the model, but read its bands as approximate, prefer conclusions that survive another sample or order (section 9), and score its forecasts out of sample (sections 10 and 11) rather than trusting its own statistics.

8 Step 5. Who leads whom? `fx.RiskGrangerλ`

p-values for every pair, cause by effect (sheet 5 Granger, by position)
`=MAKEARRAY(3, 3, LAMBDA(ci, ej, IF(ci = ej, "", INDEX(fx.RiskGrangerλ(MacroA, ModelA, ci, ej), 3, 5))))`

Cause ↓ effect →	Inflation	Unemployment	Fed funds
Inflation		0.254	0.001
Unemployment	0.015		< 0.001
Fed funds	0.001	0.121	

The clearest results are in the fed funds column: past inflation and past unemployment both help predict the rate ($p < 0.001$ and $p < 0.001$; together, $p < 0.001$). That is the Fed reacting to the economy, the Taylor rule of section 6 seen as prediction. Past unemployment helps predict inflation ($p = 0.015$), the Phillips-curve link. The rate helps predict inflation ($p = 0.001$), but in the wrong direction, as the next section shows. And the rate does *not* significantly help predict unemployment ($p = 0.121$), even though monetary policy is meant to work through jobs.

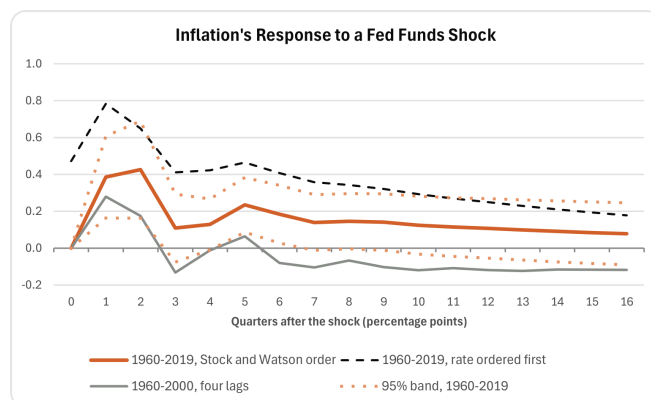
GRANGER TESTS ON PERSISTENT LEVELS

When series are close to a unit root, the usual F distribution of the Granger test is only approximate, and p-values near the 5% line deserve little weight. The Toda-Yamamoto procedure (fit one extra lag, test only the original lags) is the standard remedy. And as always, a "no" is about prediction among these three series: the rate may still matter, through channels this small model does not contain.

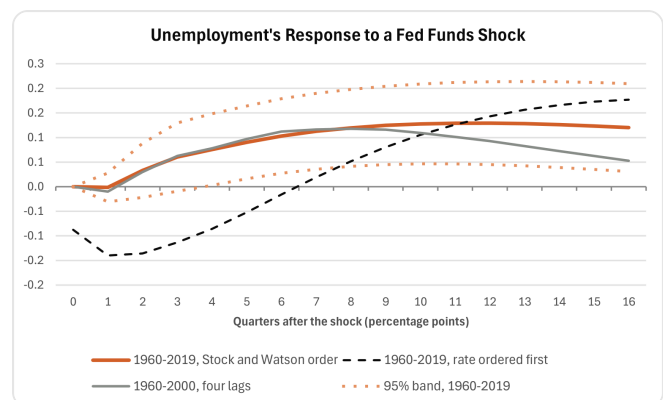
9 Step 6. A monetary-policy shock: `fx.RiskVARIRFA`, `fx.RiskVARFEVDA`

```
Model A, order infl, unemp, ff
=fx.RiskVARIRFA(ModelA, 16, "ff")
the rate first
=fx.RiskVARIRFA(ModelA, 16, "ff", , , {"ff";"infl";"unemp"})
1960-2000, 4 lags
=fx.RiskVARIRFA(fx.RiskVARFitA(TAKE(MacroA, 164), 4, , , , MacroNames), 16, "ff")
```

Stock and Watson's order puts the rate **last**: within a quarter the Fed sees inflation and unemployment and sets the rate, but they react to the rate only later. A fed funds "shock" is then the part of the rate the model cannot explain from the economy, a surprise in policy. A one-standard-deviation surprise is 0.71 points.



Inflation rises at first after a rate rise, in every version: the "price puzzle". Dotted: Model A's 95% band.



Unemployment rises slowly, as theory expects, by about a tenth of a point. Dotted: Model A's 95% band.

The price puzzle. A tighter policy should lower inflation. Instead, in Model A, inflation is 0.39 points higher a quarter after the shock and 0.43 after two, and still 0.15 after two years. This famous result appears in many VARs of this kind (Sims, 1992). The usual explanation is a missing variable. The Fed raises rates when it *expects* inflation, using information such as commodity prices that the three-variable model does

not have. The model sees rates go up and then inflation go up, and calls it a response. The puzzle is a warning about what an impulse response means: it is a causal statement only if the identifying assumptions (the order, and the shocks being truly surprises) are right.

Inflation's response to the rate shock	Quarter 1	Quarter 2	Quarter 4	Quarter 8	Quarter 12
1960-2019, order infl, unemp, ff (Model A)	0.39	0.43	0.13	0.15	0.11
1960-2019, the rate ordered first	0.78	0.65	0.42	0.34	0.25
1960-2000, four lags (Stock and Watson's sample)	0.28	0.18	-0.01	-0.07	-0.12

Robustness, the substitute for a truth. Putting the rate first lets inflation respond within the quarter (0.47 on impact), which is hard to believe economically, and changes the early quarters. On the 1960-2000 sample the puzzle fades within a year and turns negative (-0.01 at quarter 4, -0.12 at quarter 12). A conclusion that changes sign with the sample is not a conclusion. The unemployment response, a slow rise to about 0.13 points after three years, is smaller than textbook accounts but appears in every version (with the rate ordered first it dips for a year, to -0.14, before rising). Comparing versions is one way to tell a finding from noise; the bands below are the other.

Is the price puzzle real? `fx.RiskVARIRFBandsA`

```
=fx.RiskVARIRFBandsA(MacroA, ModelA, "ff", "infl", 16)
=fx.RiskVARIRFBandsA(TAKE(MacroA, 164), ModelSW, "ff", "infl", 16)    1960-2000, 4 lags
=fx.RiskVARIRFBandsA(MacroA, ModelA, "ff", "unemp", 16)
```

Quarter	Inflation, Model A		Inflation, 1960-2000		Unemployment, Model A	
	Response	95% band	Response	95% band	Response	95% band
1	0.386	0.163 to 0.608	0.280	0.067 to 0.492	-0.001	-0.031 to 0.028
2	0.426	0.163 to 0.689	0.175	-0.088 to 0.438	0.033	-0.022 to 0.088
4	0.129	-0.008 to 0.267	-0.011	-0.252 to 0.229	0.076	0.003 to 0.149
8	0.146	-0.003 to 0.296	-0.067	-0.293 to 0.160	0.120	0.041 to 0.198
12	0.108	-0.054 to 0.269	-0.118	-0.333 to 0.096	0.129	0.045 to 0.213

In this sample, yes. Model A's band for inflation lies wholly above zero in quarters 1 and 2 (0.163 to 0.608, then 0.163 to 0.689), and in 4 of the 16 quarters in all, the last of them quarter 6. The rise is not noise: from 1960 to 2019, inflation did go up after the rate surprises this model identifies. But a significant response is not a causal one. The band measures only that the coefficients are estimates. It takes the order and the three-variable model as given, and those are exactly what the missing-variable explanation doubts. On the 1960-2000 sample the first-quarter rise is significant too (0.067 to 0.492), but the later fall is not: in none of the 16 quarters does the band lie below zero. That sample cannot say that inflation falls after a tightening at all. Unemployment's slow rise, by contrast, is a finding: its band is above zero from quarter 4 to the end of the horizon.

How much do policy surprises matter?

Share of forecast-error variance, 12 quarters ahead	inflation shocks	unemployment shocks	fed funds shocks
Inflation	82%	12%	6%
Unemployment	11%	83%	6%
Fed funds rate	20%	47%	33%

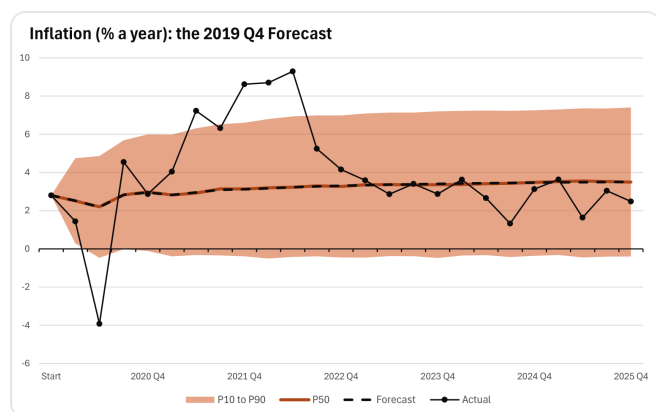
Policy *surprises* explain little of inflation's or unemployment's uncertainty: 6% and 6%. Most of the rate's own uncertainty three years out comes from the economy (47% from unemployment shocks and 20% from inflation shocks), which is what systematic policy looks like: the Fed mostly responds, it rarely surprises. That is a common finding in VARs of this kind. The decomposition depends on the order too: with the rate first, its share of inflation's variance at 12 quarters is 30%.

10 Step 7. The 2019 forecast against what happened

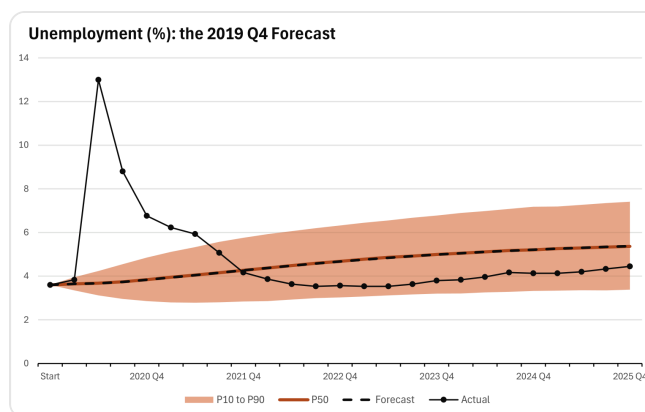
24 quarters from 2019 Q4, every series, P10 / P90 band; labels 2020 Q1 ... 2025 Q4
`=fx.RiskVARForecastλ(MacroA, ModelA, 24, , , , , QuartersAhead)`

Model A ends in 2019 Q4, so its forecast for 2020-2025 can be scored against what happened, quarter by quarter (sheet 7 *Forecast from 2019*). The *z* column measures each miss in standard errors: (actual – forecast) ÷ (the forecast's standard error). Under the model, nine quarters in ten should have *|z|* below about 1.6.

Quarter	Series	Forecast	P10	P90	Actual	<i>z</i>
2020 Q2	Unemployment	3.67	3.12	4.23	13.00	21.5
2021 Q4	Inflation	3.12	-0.39	6.64	8.62	2.0
2022 Q2	Inflation	3.23	-0.41	6.87	9.30	2.1
2023 Q4	Fed funds	3.20	-0.94	7.34	5.33	0.7



Inflation: the band held until 2021, then 4 quarters came in above its P90.



Unemployment: 2020 Q2 was 21 standard errors above the forecast.

Two kinds of miss. COVID's unemployment was not a bad forecast but an event outside the model's world: a z of 21 has no meaning under normal shocks. The inflation miss is more instructive. The band did widen as the horizon grew, but by 2022 Q2 inflation was 2.1 standard errors above the point forecast and well beyond the P90.

Right on average, wrong in runs. Over the 24 quarters, 19 inflation, 19 unemployment and 22 fed funds actuals fell inside the 80% bands, against 19.2 expected. Judged quarter by quarter the bands look fine. But the misses came in runs: unemployment above its P90 for five quarters in a row (2020 Q2 to 2021 Q2), and inflation in four of the five quarters from 2021 Q2 to 2022 Q2. Some clustering is normal: forecasts from one origin share their early errors, so misses tend to come in runs even when the model is right. But a run that starts with a 21-standard-error quarter is the signature of a break, not of bad luck. That is the signal to refit, add a dummy or rethink the model, long before the coverage count looks wrong.

The model does not know about zero. From 2021 Q1 the fed funds band's P10 is below zero, in 20 of the 24 quarters. A VAR is linear and its shocks are normal, so nothing stops a forecast rate from going negative, though the Fed has never set one. Constraints like that have to be imposed outside the VAR, for example by flooring simulated paths at zero before using them.

11 Step 8. Backtests against simple benchmarks

One forecast from one origin is one draw of luck. Sheet 8 *Backtests* refits Model A's design on the data up to each of seven origins, four years apart, forecasts the next eight quarters, and compares its RMSE with two benchmarks that need no model at all. The **random walk** says every future quarter equals the last one. **Atkeson-Ohanian** (2001) says inflation over the next year equals inflation over the last year, a benchmark that famously beat many inflation models.

the VAR's RMSE, inflation, 8 quarters after origin k (k = the origin's row in Macro)
`=INDEX(fx.RiskVARBacktestA(TAKE(Macro, k + 8), 8, LagA, , 1), 2, 2)`

Origin	Inflation			Unemployment		Fed funds	
	VAR	Random walk	Atkeson-Ohanian	VAR	Random walk	VAR	Random walk
1999 Q4	2.12	1.42	1.38	0.31	0.60	1.66	1.51
2003 Q4	1.30	2.27	1.89	0.61	0.58	1.29	1.66
2007 Q4	5.59	5.84	5.37	2.98	3.32	4.04	3.63
2011 Q4	1.91	0.94	1.97	0.65	1.02	2.26	0.06
2015 Q4	1.08	2.30	1.94	0.26	0.50	0.22	0.63
2019 Q4	3.57	3.83	4.12	4.05	4.19	2.20	1.47
2023 Q4	0.93	0.82	0.94	0.39	0.40	0.64	0.84
Average, 1999-2015	2.40	2.55	2.51	0.96	1.20	1.89	1.50

The honest scorecard. Over the five pre-COVID origins the VAR wins on unemployment (0.96 against 1.20), roughly ties on inflation with both benchmarks, and loses on the fed funds rate (1.89 against 1.50). The rate loss is mostly one origin: from 2011 Q4 the VAR expected the rate to drift back toward its long-run average, while the Fed held it near zero for years. The random walk, which just repeats the last value,

was almost perfect (0.06). Models that revert to sixty-year averages are wrong when policy deliberately stays away from them. And 2007 Q4 and 2019 Q4, the crisis origins, are bad for every method. Recessions are where forecasts matter most and where none of these does well.

Backtesting with exogenous terms. The 2023 Q4 row above was fitted with COVID left in.

`fx.RiskVARBacktestλ`'s last argument, `Exog`, takes the dummies. It fits on the rows before the holdout and uses the held-out rows of `Exog` as the future values: zeros here, because the dummies are 0 after 2020.

the 2023 Q4 origin: hold out the last 8 quarters of the whole sample, with the COVID dummies
`=INDEX(fx.RiskVARBacktestλ(Macro, 8, LagA, , 1, , , MacroNames, Covid), 2, 2)`

With the dummies the RMSEs are 0.88, 0.25 and 0.59, against 0.93, 0.39 and 0.64 without them. The dummies improve the unemployment forecast most, as section 4 suggests. Sheet *Checks* confirms the backtest equals a refit by hand: `VARFit` with `Exog`, then `VARForecast` with `FutureExog` = zeros.

12 Step 9. Model B: all the data, and 2026-2027

the dummies are 0 in every future quarter: `FutureExog`, 8 rows x 2 columns
`=fx.RiskVARForecastλ(Macro, ModelB, 8, , , , SEQUENCE(8, 2, 0, 0), NextQuarters)`
`=fx.RiskVARSimλ(Macro, ModelB, 8, "infl", , Iterations, 1, SimSeed, SEQUENCE(8, 2, 0, 0))`

Model B (264 quarters, the COVID dummies) is the model to use going forward. A model with exogenous columns needs their future values to forecast; here that is simple: no more pandemic quarters, so zeros. With a real outside driver (an oil price, a policy path), `FutureExog` is where the scenario goes. Model B's largest root is 0.962 and its residuals still fail the portmanteau test (p 7.4e-06). Six more years also change the Granger results. The rate's effect on unemployment comes closer to significance (p 0.080, from 0.121), while unemployment's effect on inflation weakens (p 0.114, from 0.015): the Phillips-curve link of 1960-2019 did not survive the 2020s.

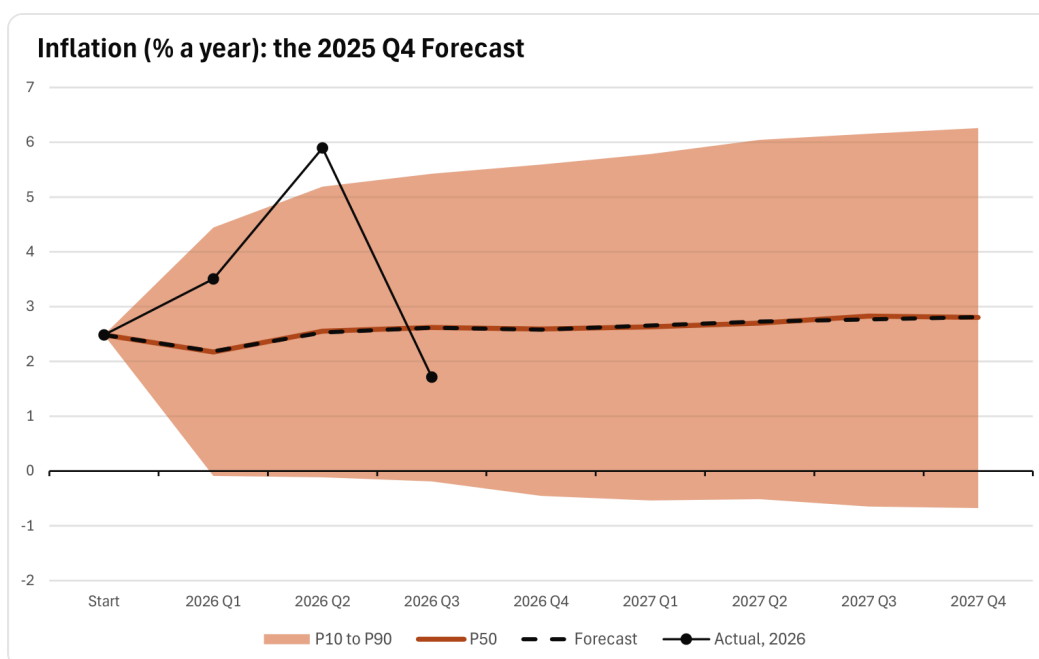
Model B's forecast from 2025 Q4	2026 Q1	2026 Q4	2027 Q4	2027 Q4: P10	P90
Inflation (% a year)	2.18	2.58	2.81	-0.66	6.28
Unemployment (%)	4.53	4.68	4.79	3.48	6.11
Fed funds rate (%)	3.81	3.84	3.88	0.70	7.05

SINCE THE DATA ENDED

The workbook stops the model at 2025 Q4, but FRED already has the first three quarters of 2026. Against the forecast:

2026	Inflation: actual	forecast	P10 to P90	Unemployment: actual	forecast	P10 to P90
Q1	3.51	2.18	-0.08 to 4.45	4.33	4.53	4.23 to 4.84
Q2	5.89	2.53	-0.12 to 5.17	4.27	4.59	4.04 to 5.14
Q3	1.71	2.62	-0.19 to 5.42	4.13	4.64	3.88 to 5.40

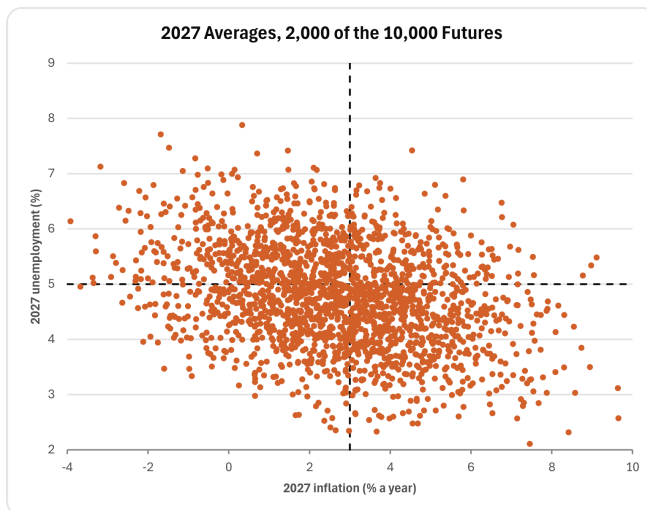
Unemployment has stayed a little below the forecast and inside its band every quarter. Inflation has been noisier and higher: its three quarters average 3.7% against forecasts averaging 2.4%, and the second quarter came in above its P90. One quarter outside an 80% band is expected one time in five; if the next ones follow it, that is the run of section 10 again.



Inflation, 2026-2027: Model B's band, its forecast (dashed) and the three quarters of 2026 so far (black).

Joint questions need joint paths

A planner may care less about either series alone than about a combination: what is the chance that 2027 brings both inflation above 3% and unemployment above 5% (stagflation), with both thresholds as yellow cells? Each `fx.RiskVARSimλ` call returns one series, but with the same VarID and Seed the three sets of paths are the same 10,000 futures (Part 1, section 12), so the question is one formula: `=AVERAGE((Infl27 > InflLine) * (Unemp27 > UnempLine))`.



2027 averages, 2,000 of the 10,000 futures, with the two thresholds. The top-right corner is the stagflation outcome.

2027 (averages of the four quarters)	Chance
Inflation above 3%	45%
Unemployment above 5%	39%
Both, from the joint paths	13%
Both, if they were independent	18%
Fed funds rate below zero in 2027 Q4	6%

Inflation and unemployment are negatively correlated in the paths (-0.29), so the joint outcome is less likely than multiplying the two chances suggests. And 6% of the futures have a negative rate in 2027 Q4: the zero bound again, which a planner would floor before using the paths.

13 Lessons from real data

- **Write down the data choices.** A filled month, a seasonal adjustment, a revision: each changes the numbers. Keep the raw data next to the used data, and the fill in a visible cell.
- **Choose the sample on purpose.** One extraordinary event can dominate a least-squares fit. Stop before it, or give it a dummy and say so. Do not let it decide the coefficients silently.
- **Measure fit against the right yardstick.** R^2 in levels flatters persistent series; judge forecasts by their errors out of sample.
- **Expect the diagnostics to fail, and read the bands accordingly.** Real residuals have fat tails and changing volatility; the 80% band is a guide, not a guarantee.
- **Robustness replaces the truth.** Report impulse responses with their bands, under more than one order and sample. A response can be significant and still misidentified (the price puzzle): it is a question, not an answer.
- **Always compare with a benchmark.** A random walk and a four-quarter average are hard to beat for inflation and interest rates. The VAR earns its place where the series really do lead one another, here unemployment.
- **A linear model knows no floors.** Interest rates at zero, prices that cannot go negative: impose them on the simulated paths before using them.
- **Use the joint paths for joint questions.** Same VarID and Seed for every series, and the probabilities of combinations follow from one formula.

14 Checking the model, and function notes

Sheet *Checks* lists 15 tests with value, expected and difference, and a blank *Manual sign-off* column for your ✓. Among them: each quarterly CPI is the average of its months (with the October fill); Model A's 2020 Q1 forecast equals its coefficients times the last values; a fed funds shock moves inflation only from the next quarter, because the rate is ordered last; the pulse dummies leave zero residuals in 2020 Q2 and

Q3 (8.9e−15); the band around inflation's response is the response $\pm$ 1.96 standard errors; the 2023 backtest with `Exog` equals a refit by hand; and the simulated paths' means and correlation match the formulas within three standard errors.

What this guide adds to Part 1's function reference

Argument	Functions	Meaning
<code>Exog</code>	<code>VARSelect</code> , <code>VARFit</code> , <code>VARSummary</code> , <code>VARResids</code> , <code>Granger</code> , <code>VARPortmanteau</code> , <code>VARIRFBands</code> , <code>VARBacktest</code>	$n \times e$ exogenous columns, the same rows as Data (dummies, outside drivers). Pass the same block to every function that reads a model fitted with them.
<code>Names</code>	<code>VARFit</code>	The series' names, then the exogenous columns' names, e.g. <code>HSTACK(MacroNames, {"covid_q2", "covid_q3"})</code> .
<code>FutureExog</code>	<code>VARForecast</code> , <code>VARSim</code>	Horizon $\times$ e future values of the exogenous columns: zeros for pulse dummies, a scenario path for a driver.
<code>Order</code>	<code>VARIRF</code> , <code>VARIRFBands</code> , <code>VARFEVD</code>	The Cholesky order as names, e.g. <code>{"ff"; "infl"; "unemp"};</code> the default is the column order.
<code>Level</code>	<code>VARIRFBands</code>	The band's confidence level: 0.95 by default; 0.68 gives the one-standard-error bands many papers show.

KNOWN LIMITS

`fx.RiskVARBacktest $\lambda$` uses one origin per call (call it on `TAKE(Data, k + h)` for others). The bands of `fx.RiskVARIRFBands $\lambda$` are asymptotic: they take the model and the order as given, and with a root near 1 (Model A's is 0.955) they can be unreliable at long horizons, where a bootstrap is the usual remedy. XL Edge has no bootstrap bands.

References: Sims (1980), "Macroeconomics and Reality", *Econometrica* 48(1); Sims (1992), "Interpreting the Macroeconomic Time Series Facts", *European Economic Review* 36(5); Sims, Stock and Watson (1990), *Econometrica* 58(1); Stock and Watson (2001), "Vector Autoregressions", *Journal of Economic Perspectives* 15(4); Atkeson and Ohanian (2001), *FRB Minneapolis Quarterly Review* 25(1); Toda and Yamamoto (1995), *Journal of Econometrics* 66; Lütkepohl (2005), *New Introduction to Multiple Time Series Analysis*. Data: U.S. Bureau of Labor Statistics and Board of Governors of the Federal Reserve System, retrieved from FRED, Federal Reserve Bank of St. Louis. The VAR functions are validated against statsmodels. XL Edge and the `fx.Risk...` functions: © 2026 W Phillips, edgewisedata.com, MIT License.