

Vector autoregression: forecasting revenue from the orders pipeline

Industry orders lead a manufacturer's bookings, and bookings lead its revenue. A vector autoregression (VAR) turns that chain into a forecast with honest bands. Here it gives the 95-million-dollar FY2027 revenue budget a 36% chance, and values a one-standard-deviation surprise in industry orders at about 3.1 million dollars of next year's revenue.

Functions `fx.RiskVARSelectλ` · `VARFitλ` · `VARSummaryλ` · `VARResidsλ` · `VARStabilityλ` · `VARPortmanteauλ` · `Grangerλ` · `VARIRFλ` · `VARIRFBandsλ` · `VARFEVDλ` · `VARForecastλ` · `VARSimλ` · `VARBacktestλ`

In XL Edge XL Edge tab > Productivity Tools > Stat Tools > Insert Time Series > Multivariate - VAR and cointegration

Example *XL Edge Example - VAR 1 Business Forecasting.xlsx* (same folder). Its history is simulated from a model we know, so every estimate in this guide can be graded against the truth. Every number here is read from the workbook.

Series Part 1 of 2. Part 2, *VAR 2 Real Macro Data*, fits the same functions to US inflation, unemployment and interest rates, where nobody knows the true model.

Needs Excel for Microsoft 365. The functions also work without XL Edge once they are in a workbook.

1 The question, and why a VAR

2 What a VAR is, in plain English

3 The workflow and the ribbon

4 Step 0. Growth rates, not levels

5 Step 1. How many lags?

6 Step 2. Fit the VAR and read it

7 Step 3. Is the model sound?

8 Step 4. Who leads whom?

9 Step 5. How shocks travel

10 Step 6. Forecast and fan chart

11 Step 7. Backtest honestly

12 Step 8. From paths to the budget

13 Under the hood

14 Checking the model

15 Pitfalls and good practice

16 Function reference

1 The question, and why a VAR

The manufacturer from the Copula and Plant 2 guides closed September 2026 with 82.0 million dollars of revenue over the last twelve months. Finance is setting the FY2027 budget, and the sales plan says 95 million. Two questions come up every year:

- **How likely is the budget?** Not "is it achievable", but what are the odds, and what range of outcomes should the plan prepare for?
- **What should we watch?** Revenue ships from bookings, and bookings follow the industry's new orders. If industry orders surprise this month, what does that do to next year?

Three monthly series hold the answer: **industry** new orders (a published indicator), the company's **bookings** (orders received) and its **revenue** (orders shipped). They feed each other with delays, and their surprises arrive together. A forecast of revenue alone would ignore the two series that already contain part of its future. A regression of revenue on bookings would need a forecast of bookings first, and bookings in turn depend on industry orders. A **vector autoregression** forecasts all three at once, from their joint past, and keeps the correlation between their surprises.

WHY SIMULATED DATA

The workbook's ten years of history were generated from a VAR whose coefficients we chose (sheet *True model*). In practice the true model is never known; here it is, so each step can show not only what the function estimates but how close that is to the truth. Part 2 does the same analysis on real data, where all you have are the estimates and their standard errors.

2 What a VAR is, in plain English

A VAR with p lags explains **each** series by the last p values of **every** series, plus a constant. With two lags, revenue's equation reads:

$$\begin{aligned} \text{revenue}(t) = & c + a_1 \text{industry}(t-1) + b_1 \text{bookings}(t-1) + d_1 \text{revenue}(t-1) \\ & + a_2 \text{industry}(t-2) + b_2 \text{bookings}(t-2) + d_2 \text{revenue}(t-2) + e(t) \end{aligned}$$

Industry and bookings get equations of the same form, with their own coefficients. Each equation is an ordinary least-squares regression on the same lagged values (section 13 rebuilds one with `LINEST`). What makes it a *system* is the shocks $e(t)$: the three equations' surprises are correlated, and that correlation, the **residual covariance** Σ , is part of the model.

A worked example. The fitted revenue equation forecasts October 2026's revenue growth at **2.54%**. Most of that comes from bookings: they grew 2.18% in September and 1.85% in August, with coefficients 0.53 and 0.70. Orders booked in the last two months are still shipping. Revenue's own past enters with *negative* signs (-0.36 and -0.20). That does not mean revenue falls after a good month. It means part of a good month is timing: an order shipped early in one month is not there to ship the next, so a strong month is partly paid back.

The vocabulary, once

Term	Meaning
Lag	A past value: bookings(t-1) is last month's bookings growth. A VAR(2) uses two lags of every series.
Stationary	Keeps returning to an average. Growth rates usually are; levels (revenue in dollars) usually are not.
Model block	The table <code>fx.RiskVARFit</code> writes: settings, fit statistics, coefficients and Σ . Every later function reads it.
Residual covariance Σ	How large each equation's surprises are, and how they move together in the same month.
Roots (stability)	Numbers that say how fast shocks die out. All below 1 in size: the model settles down.
Granger causality	X's past helps predict Y beyond Y's own past. A statement about prediction, not cause.
Orthogonalized shock	A surprise in one series, separated from the others' same-month surprises by a Cholesky factor of Σ .
Cholesky order	Who may move whom within the month: earlier series move later ones, not the reverse. An assumption.
Impulse response (IRF)	How each series responds, month by month, to a one-standard-deviation shock.
Band (of a response)	The range a response could plausibly take, because the coefficients behind it are estimates. Where it contains 0, the data cannot tell the response from none.
Variance decomposition (FEVD)	The share of a series' forecast uncertainty that each shock explains.
Portmanteau test	Is any correlation left in the residuals? If so, the model has missed a pattern.

3 The workflow and the ribbon

Step	Question	Function	Sheet
0	Do the series wander, or return to an average?	<code>fx.RiskUnitRootλ</code>	0 Prepare
1	How many lags?	<code>fx.RiskVARSelectλ</code>	1 Lags
2	Fit, then read the coefficients	<code>fx.RiskVARFitλ</code> , <code>fx.RiskVARSummaryλ</code>	2 Fit
3	Anything left in the residuals? Stable?	<code>fx.RiskVARResidsλ</code> , <code>fx.RiskVARPortmanteauλ</code> , <code>fx.RiskVARStabilityλ</code>	3 Diagnostics
4	Who leads whom?	<code>fx.RiskGrangerλ</code>	4 Granger
5	How does a shock spread?	<code>fx.RiskVARIRFλ</code> , <code>fx.RiskVARIRFBandsλ</code> , <code>fx.RiskVARFEVDλ</code>	5 Shocks
6	Where are they heading?	<code>fx.RiskVARForecastλ</code> , <code>fx.RiskVARSimλ</code>	6 Forecast
7	Would it have worked before?	<code>fx.RiskVARBacktestλ</code>	7 Backtest
8	What are the budget's odds?	<code>fx.RiskVARSimλ</code> paths, Excel	8 Budget

Everything after step 2 reads the **Model block**, so you fit once and point every later function at it. The example names it `Model` (Formulas > Name Manager), along with `GrowthRates`, `SeriesNames` and `LagOrder`, so the formulas below read like sentences. Change the data, and the whole chain recalculates.

XL Edge tab > **Productivity Tools** > **Stat Tools (arrow)** > **Insert Time Series** > **Multivariate - VAR and cointegration**

Menu items	XL Edge asks for
Select VAR Lags, Fit VAR (Model Block)	The data: one column per series, oldest row first, names in the header row above. Then the maximum lags, or the lags.
VAR Coefficient Summary, VAR Residuals, Granger Causality, Residual Portmanteau Test, VAR Forecast, VAR Simulated Paths	The data and the Model block (its top-left cell, or the whole block), then any options.
VAR Stability, Impulse Responses, Variance Decomposition	The Model block only.
VAR Backtest	The data, the holdout length, the lags and the series to score.

Each item tells you the size of its result, writes one live formula, and installs the functions the workbook needs. *VAR Simulated Paths* feeds *Insert Monte Carlo Chart* > *Fan Chart (Time Series)*.

4 Step 0. Growth rates, not levels: `fx.RiskUnitRootλ`

Sheet *Data* holds the monthly levels from September 2016: industry orders in millions of dollars, bookings and revenue in thousands (revenue 4,397 then, 7,092 in September 2026). A VAR needs **stationary** series, ones that keep returning to an average. Levels that trend and wander have no average to return to, and regressions between wandering series find relationships that are not there (a *spurious regression*). So test the levels and the growth rates:

```
=fx.RiskUnitRootλ(100 * LN(Data!$F$10:$F$130), "ct", , , "log revenue")
=fx.RiskUnitRootλ(CHOOSECOLS(GrowthRates, 3), , , , "revenue growth")
```

Test (what it assumes until rejected)	log revenue: p	verdict	revenue growth: p	verdict
ADF (a unit root)	0.453	Unit root	0.005	Stationary
Phillips-Perron (the same)	0.649	Unit root	< 0.001	Stationary
KPSS (stationary)	<0.01	Not stationary	>0.10	Stationary

All three agree, and industry orders tell the same story (sheet *O Prepare*): the levels wander, the growth rates do not. So the VAR models **growth**: $100 \times \text{LN}(\text{this month} \div \text{last month})$, about the percentage change. The growth rates are on *Data* in H11:J130, named **GrowthRates**, with their names in H9:J9 (**SeriesNames**).

ONE THING GROWTH RATES LEAVE OUT

Bookings and revenue cannot drift apart for long: what is booked gets shipped. Their twelve-month book-to-bill ratio stayed between 0.92 and 1.07 for ten years. A VAR in growth rates does not know about that tie. Over the fifteen months this guide forecasts, that does little harm, but for a long-range plan the tie matters, and a *vector error-correction model* keeps it (see the *Multivariate Time Series* guide).

5 Step 1. How many lags? **fx.RiskVARSelectλ**

```
=fx.RiskVARSelectλ(GrowthRates, 12)
```

This fits every order from 1 to 12 lags on the same 108 months, so the comparison is fair, and scores each with four information criteria. Each criterion is the fit (the log of the residual covariance's determinant) plus a penalty per coefficient, and the lowest score wins:

Lags	AIC	HQ	SC (BIC)	FPE	Coefficients per equation
1	4.550	4.671	4.848	94.7	4
2	3.210	3.421	3.731	24.8	7
3	3.227	3.529	3.972	25.2	10
12	3.754	4.872	6.511	46.6	37

All four choose **2 lags**, which is the true model's order. Going from one lag to two improves the fit by far more than the penalty, because the backlog ships over two months. A third lag adds three coefficients per equation and nothing to the fit, so every criterion gets worse. AIC and FPE penalize least and lean toward longer lags; SC (Schwarz, also called BIC) penalizes most and leans shorter; HQ sits between. When they disagree, which is common with real data (Part 2), fit the candidates and let the diagnostics and the backtest decide. The example stores the AIC's choice in the name **LagOrder**, and the fit reads it from there.

6 Step 2. Fit the VAR and read it: `fx.RiskVARFit`, `fx.RiskVARSummary`

```
data, lags, Type (const by default), Season, Exog, Names
=fx.RiskVARFit(GrowthRates, LagOrder, , , , SeriesNames) '2 Fit'!C10, named Model
=fx.RiskVARSummary(GrowthRates, Model)
```

The **Model block** is 25 rows × 4 here: (15 + m + K) × (K + 1), with K = 3 series and m = 7 terms per equation (3 × 2 lags + a constant). From the top:

- rows 2-8, the specification: Variables, Lags, Type, Season, Exogenous, Observations (118), Sample (120);
- rows 9-13, the fit statistics: LogLik, AIC, HQ, SC, FPE;
- row 14 names the equations; then one row per term (`industry.l1` ... `revenue.l2`, `const`), one column per equation;
- the last four rows, Σ .

It is plain cells. Sheet *True model* is the same layout typed in by hand, and every VAR function accepts it.

The **summary** gives each equation's coefficients with standard errors, t-statistics and p-values, then its R^2 , and the residual covariance and correlation matrices. The R^2 values are 0.40 (industry), 0.28 (bookings) and **0.79** (revenue). Revenue is the most predictable of the three because most of its next month is already in the order book.

The residual correlations say which surprises arrive together: bookings and revenue 0.44, industry and bookings 0.27, industry and revenue only 0.05.

The revenue equation (sheet *2 Fit*), against the true model

Term	True value	Estimate	Std error	p-value	Reading
<code>bookings.l1</code>	0.50	0.53	0.05	< 0.001	last month's orders ship this month
<code>bookings.l2</code>	0.70	0.70	0.05	< 0.001	... and so do the month before's
<code>revenue.l1</code>	-0.30	-0.36	0.05	< 0.001	timing: a strong month is partly paid back
<code>revenue.l2</code>	-0.20	-0.20	0.05	< 0.001	
<code>industry.l1</code>	0.00	-0.08	0.09	0.410	no direct effect: it works through bookings
<code>industry.l2</code>	0.00	0.15	0.10	0.135	a true zero, estimated at 0.15
<code>const</code>	0.12	0.14	0.13	0.284	

"NOT SIGNIFICANT" MEANS "UNKNOWN", NOT "ZERO"

The true effect of industry orders two months back is zero, and the estimate is 0.15 with a standard error of 0.10 (p 0.135). A rough 95% range is the estimate ± 2 standard errors: -0.05 to 0.35. That range contains zero, so the test cannot reject zero, but it also contains effects as large as 0.35. A high p-value says the data cannot tell. It does not say the effect is absent. Dropping a term because it is "not significant" is a judgment, not a finding.

Graded against the truth. The block on the right of sheet *2 Fit* divides each of the 21 estimates' distance from its true value by its standard error. About 95% should lie within 2. Here **21 of 21** do, the largest at 1.51. The standard errors are honest: they describe how far off the estimates really are.

7 Step 3. Is the model sound?

Anything left in the residuals? `fx.RiskVARResiduals`, `fx.RiskVARPortmanteau`

```
=fx.RiskVARResiduals(GrowthRates, Model)          sheet Residuals: actual less fitted  
=fx.RiskVARPortmanteau(GrowthRates, Model, 10)
```

If the model has captured the dynamics, its residuals are *white noise*: no correlation with their own past or with each other's past. The multivariate portmanteau test checks every auto- and cross-correlation up to 10 months back at once. Comparing the right model with one that has too few lags shows what it catches:

Model	Adjusted statistic	Degrees of freedom	p-value	Verdict
VAR(2), the fit	77.0	72	0.321	White noise
VAR(1): one lag too few	197.4	81	1.0e-11	Serial correlation

With one lag, the bookings from two months back have nowhere to go but into the residuals, and the test sees it plainly. Underfitting always shows up this way. The degrees of freedom are $K^2 \times (10 - p)$, so the test needs more lags than the model. For a single residual column, `fx.RiskACF` on the *Residuals* sheet shows which lag the leftover correlation sits at. `fx.RiskNormality` on the same sheet finds the residuals normal (Jarque-Bera p 0.97 for revenue). That is expected, since the shocks were simulated as normal. Real data rarely are (Part 2).

Stability: `fx.RiskVARStability`

```
=fx.RiskVARStability(Model)
```

A VAR is **stable** when every root of its companion matrix (section 13) has a modulus below 1. Then shocks die out and forecasts settle at the long-run averages. A VAR(2) in three series has 6 roots. The verdict is **Stable**, with largest modulus 0.769 (the true model's: 0.739).

What the number means. Each month, the echo of a shock shrinks by roughly the largest modulus, so it halves in $\text{LN}(0.5) \div \text{LN}(0.769) = 2.6$ months. A modulus near 1 means shocks linger for years. That happens when levels that wander are modeled without differencing, or when there are too many lags for the data.

8 Step 4. Who leads whom? `fx.RiskGranger`

```
does X's past help predict Y, beyond Y's own past and the other series?  
=fx.RiskGranger(GrowthRates, Model, "industry", "bookings")  
=fx.RiskGranger(GrowthRates, Model, {"bookings","revenue"}, "industry")
```

Cause → effect	F	p-value	Verdict (5%)	Same month: p	True model
industry → bookings	19.17	1.3e-08	Granger-causes	0.004	yes
bookings → revenue	130.05	1.8e-42	Granger-causes	< 0.001	yes
industry → revenue	1.17	0.311	No Granger causality	0.60	no (only via bookings)
revenue → industry	0.68	0.505	No Granger causality	0.60	no
bookings and revenue → industry	0.38	0.822	No Granger causality	0.012	no

The tests recover the true structure. Industry orders help predict bookings, and bookings help predict revenue. Industry orders do *not* help predict revenue once bookings are in the model (p 0.311). Their effect already runs through the order book, so the order book is the indicator to watch. And a company this size does not move its industry: the last two rows find no Granger causality, which is the truth. The F test checks that all the cause's lags in the effect's equation are zero, with 2 and 333 degrees of freedom. The "same month" column tests *instantaneous* causality: correlated surprises within the month, which a VAR cannot order by itself (next section).

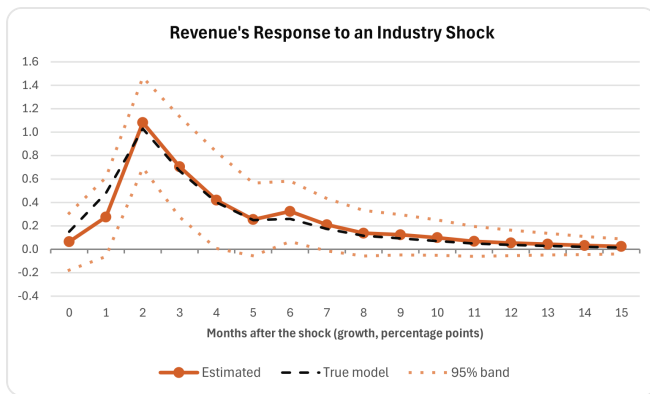
PREDICTION, NOT CAUSE

"Bookings Granger-cause revenue" means past bookings improve revenue forecasts. That matches the business here, but the test cannot know it. A series that merely reacts faster to a common cause also "Granger-causes" the slower one, and anything left out of the model can create or hide the pattern.

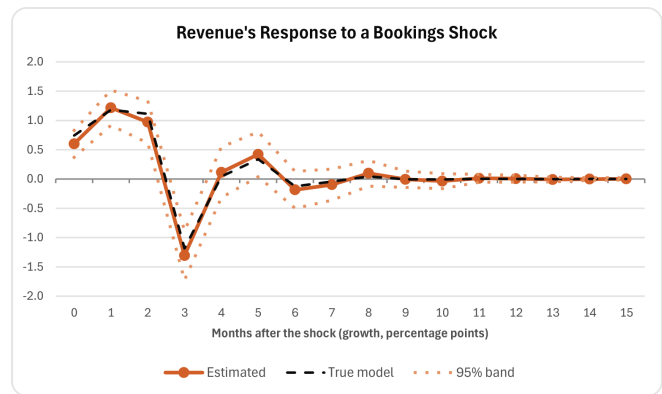
9 Step 5. How shocks travel: `fx.RiskVARIRFA`, `fx.RiskVARFEVD`

```
all 9 shock → response pairs, months 0 to 15 (orthogonalized, Cholesky in column order)
=fx.RiskVARIRFA(Model, 15)
one shock, added up: for growth rates, the effect on the level
=fx.RiskVARIRFA(Model, 15, "industry", , TRUE)
one shock → response pair, with its 95% band
=fx.RiskVARIRFBands(GrowthRates, Model, "industry", "revenue", 15)
```

An **impulse response** asks: if one series surprises by one standard deviation this month, how do all three respond over the following months? Because same-month surprises are correlated, they are first separated with a Cholesky factor of Σ , in an **order**. Series earlier in the order may move later ones within the month, not the reverse. The example's column order, industry, bookings, revenue, matches the business: one company cannot move its industry within a month, and this month's orders can ship this month, but this month's shipments cannot create this month's orders.



An industry surprise (1.39% in a month) reaches revenue through bookings and peaks 2 months later. Dotted: the 95% band.



A bookings surprise (2.71%) ships over three months, then is partly paid back. Dotted: the 95% band.

Reading the charts. After an industry shock, revenue growth rises by 0.06, 0.28 and then 1.08 points in months 0, 1 and 2. The true model says 0.15, 0.48 and 1.03. After a bookings shock, revenue rises for three months (0.60, 1.22, 0.97), then dips (−1.31 in month 3): shipments pulled forward leave a gap. The estimated responses track the true ones closely, because 120 months pin these coefficients down well.

How sure are the responses? `fx.RiskVARIRFBandsA`

Each response is built from estimated coefficients, so it has a standard error of its own.

`fx.RiskVARIRFBandsA` returns one shock → response pair, month by month: the response, the lower and upper limits of a band (95% unless you pass `Level`) and the standard error. Those limits are the dotted lines on the charts. Where the band contains zero, the data cannot tell that month's response from no response at all.

Revenue's response (growth, points)	Month 0	Month 1	Month 2	Month 3
To an industry shock: estimate	0.06	0.28	1.08	0.71
95% band	−0.18 to 0.31	−0.06 to 0.61	0.70 to 1.47	0.28 to 1.13
True model	0.15	0.48	1.03	0.67
To a bookings shock: estimate	0.60	1.22	0.97	−1.31
95% band	0.37 to 0.83	0.92 to 1.52	0.61 to 1.33	−1.73 to −0.89
True model	0.74	1.18	1.11	−1.18

The industry band contains zero until month 2. Yet the true response in month 1 is 0.48, a real effect: 120 months of data cannot tell it from zero. A band that contains zero means "unknown", not "none", just as for a coefficient in section 6. In month 2, when the extra orders ship, the band is well clear of zero (0.70 to 1.47). The bookings band is clear of zero in each of months 0 to 3, the dip in month 3 included (−1.73 to −0.89): the payback is real, not noise. Of the 16 true responses, 16 (industry) and 16 (bookings) lie inside their bands; a 95% band should miss only about one month in twenty. Sheet 5 *Shocks* lists every month with the true response beside it.

WHAT THE BAND COVERS, AND WHAT IT LEAVES OUT

The band measures one uncertainty: the coefficients (and, for one-standard-deviation shocks, Σ) are estimates. It is the mirror image of the forecast band in section 10, which covers future surprises but treats the coefficients as known. It takes the model and the Cholesky order as given, so test those another way: another order, another lag length, a shorter sample (Part 2 does this). And it is asymptotic, Lütkepohl's delta method, the standard errors statsmodels reports. With a short sample or a root near 1, a bootstrap gives more reliable bands. Papers often show 68% bands, one standard error either side: pass

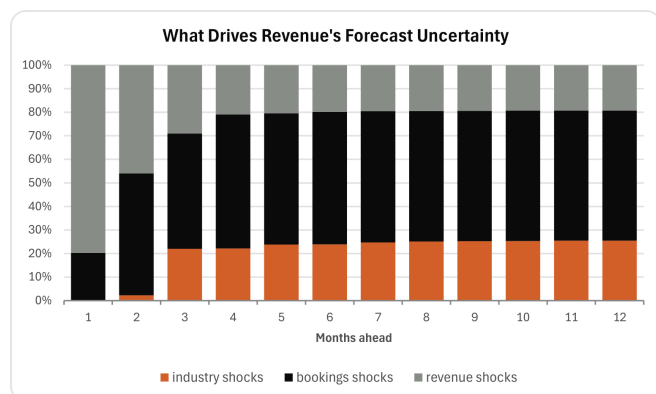
Level 0.68.

From a growth response to dollars

For growth rates, the **Cumulative** option adds the responses up, which gives the effect on the *level*. Fifteen months after an industry surprise, revenue is 3.92% higher than it would have been (the true model: 3.84%). Applied month by month to the point forecast of FY2027 revenue (sheet 5 *Shocks*), a one-standard-deviation industry surprise in October 2026 is worth **3.1 million dollars** of FY2027 revenue (true model: 3.2 million). That is the early-warning value of the indicator: an industry print well above or below expectations moves next year's revenue by about that much, before a single order arrives.

What drives the forecast's uncertainty

```
=fx.RiskVARFEVDA(Model, 12, "revenue")
=fx.RiskVARFEVDA(Model, 12, "revenue", {"revenue";"bookings";"industry"})    order reversed
```



Share of revenue's forecast-error variance due to each shock, 1 to 12 months ahead.

Months ahead	industry	bookings	revenue	revenue, order reversed
1	0%	20%	80%	100%
3	22%	49%	29%	35%
12	25%	55%	19%	26%

Shares of revenue's forecast-error variance by shock; the last column puts revenue first in the Cholesky order.

One month ahead, revenue's uncertainty is mostly its own: shipping surprises. A year ahead, more than half comes from bookings and a quarter from the industry. To narrow next year's revenue range, the order book matters far more than the shipping schedule. Reverse the order, putting revenue first, and its own share at 12 months rises from 19% to 26%. In month one it becomes all of it, because revenue now absorbs the shared same-month surprise. The conclusion survives the change, but the exact shares depend on the order.

TRY THIS

On sheet 5 *Shocks*, add an **Order** to the IRF, for example {"revenue";"bookings";"industry"}, and watch the month-0 responses change while the later months barely move. Then change **Horizon** to 24.

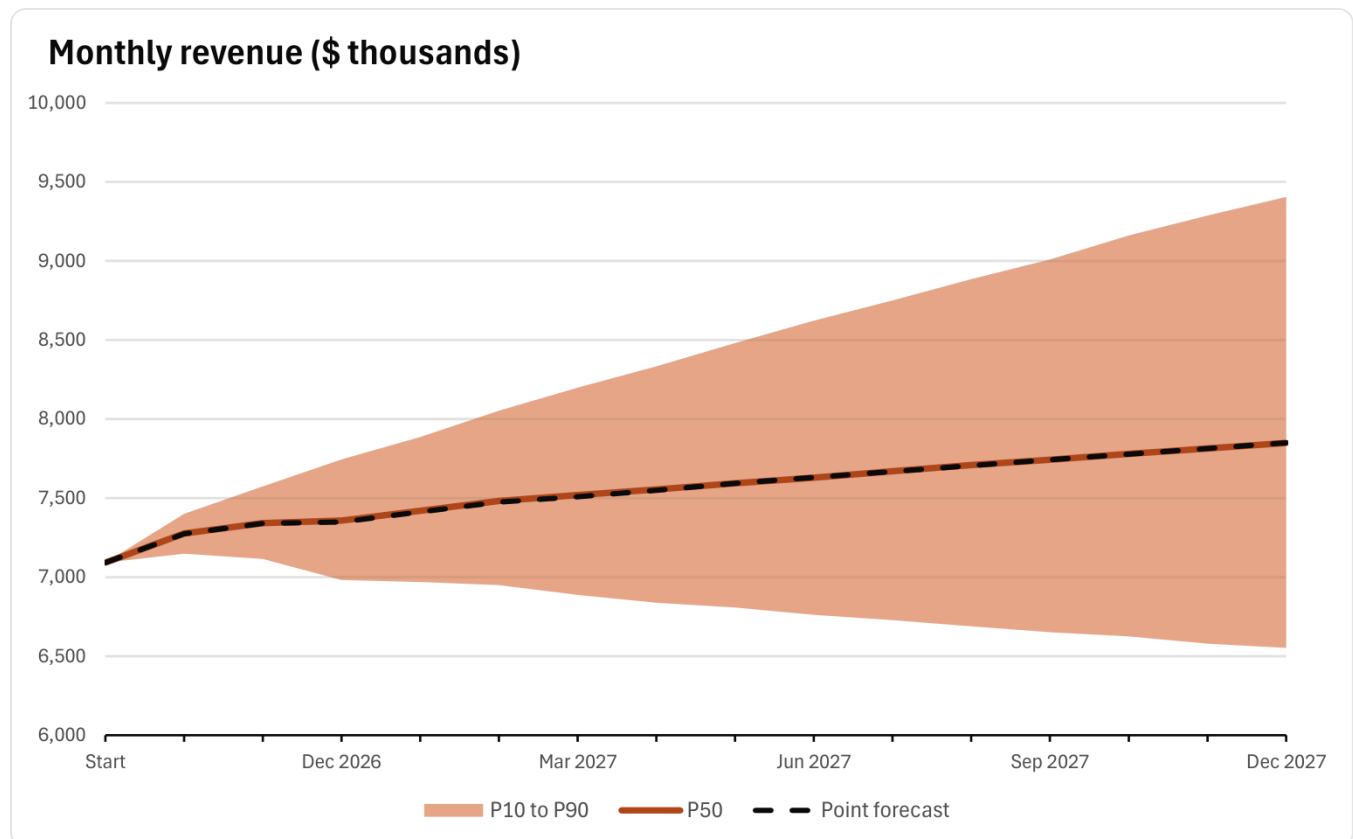
10 Step 6. Forecast and fan chart: `fx.RiskVARForecastλ`, `fx.RiskVARSimλ`

```
15 months, every series, P10 / P90 band; labels Oct 2026 ... Dec 2027
=fx.RiskVARForecastλ(GrowthRates, Model, 15, , , , FcLabels)
10,000 simulated 15-month paths of revenue growth (Iterations, SimSeed: yellow cells)
=fx.RiskVARSimλ(GrowthRates, Model, 15, "revenue", , Iterations, 1, SimSeed)
```

The forecast table gives, for each series and month, the point forecast and a P10-P90 band from the model's forecast-error variance (pass `Lower` and `Upper` for other bands). Revenue growth:

Month	Forecast (%)	P10	P90	Why
October 2026	2.54	0.81	4.26	August and September bookings still shipping
November 2026	0.89	-1.53	3.31	part of October is paid back
December 2026	0.14	-2.91	3.20	
December 2027	0.46	-3.30	4.21	back to the long-run average

The forecasts settle toward the long-run growth the coefficients imply, 0.44% a month for revenue (sheet *6 Forecast*; the truth is 0.40). The band widens for a few months and then stops widening, as a stable model's must. One month ahead, the band is just the forecast $\pm 1.28 \times \sqrt{\Sigma}$ (section 13). Further out, the echoes of earlier shocks add up.



XL Edge's fan chart of the 10,000 simulated paths, turned into dollars: P10-P90 band, median (solid) and the point forecast (dashed, passed to `fx.RiskChartFanλ` as its Plan). The value axis was zoomed by hand (Format Axis), because Excel's automatic axis starts at 0.

From growth to dollars. Each path of growth rates becomes a path of revenue: `RevNow × EXP(running total of growth ÷ 100)`, done for all 10,000 paths at once with one `MMULT` (sheet *Paths (revenue)*). Monthly revenue's P10-P90 range widens from 7,149-7,401 thousand in October to 6,553-9,405 thousand by December 2027. The levels keep spreading after the growth band has stopped widening, because each month's surprise is carried into every later month's level.

`fx.RiskVARSimA` draws correlated shocks for **all three** series and lets them feed through the model. Their percentiles reproduce the forecast band, checked on sheet *Checks*. Called once per series with the **same VarID and Seed**, the paths belong together: row *r* of the revenue paths and row *r* of the bookings paths are the same simulated future (section 12 shows why that matters).

THE BANDS LEAVE OUT PARAMETER UNCERTAINTY

The band treats the estimated coefficients as if they were the truth, as vars and statsmodels do by default, so it is a little too narrow, more so for short samples. The backtest in the next section is the honest check.

11 Step 7. Backtest honestly: `fx.RiskVARBacktestA`

```
=fx.RiskVARBacktestA(GrowthRates, 12, LagOrder, , "revenue", , , , SeriesNames)
```

The backtest refits the VAR without the last 12 months (October 2025 to September 2026), forecasts them from September 2025, and scores revenue growth against what happened. A forecast is only useful against an alternative, so sheet *7 Backtest* adds two simple ones. A naive forecast uses the average growth before the holdout (0.35% a month). A "same as last month" forecast repeats the last value.

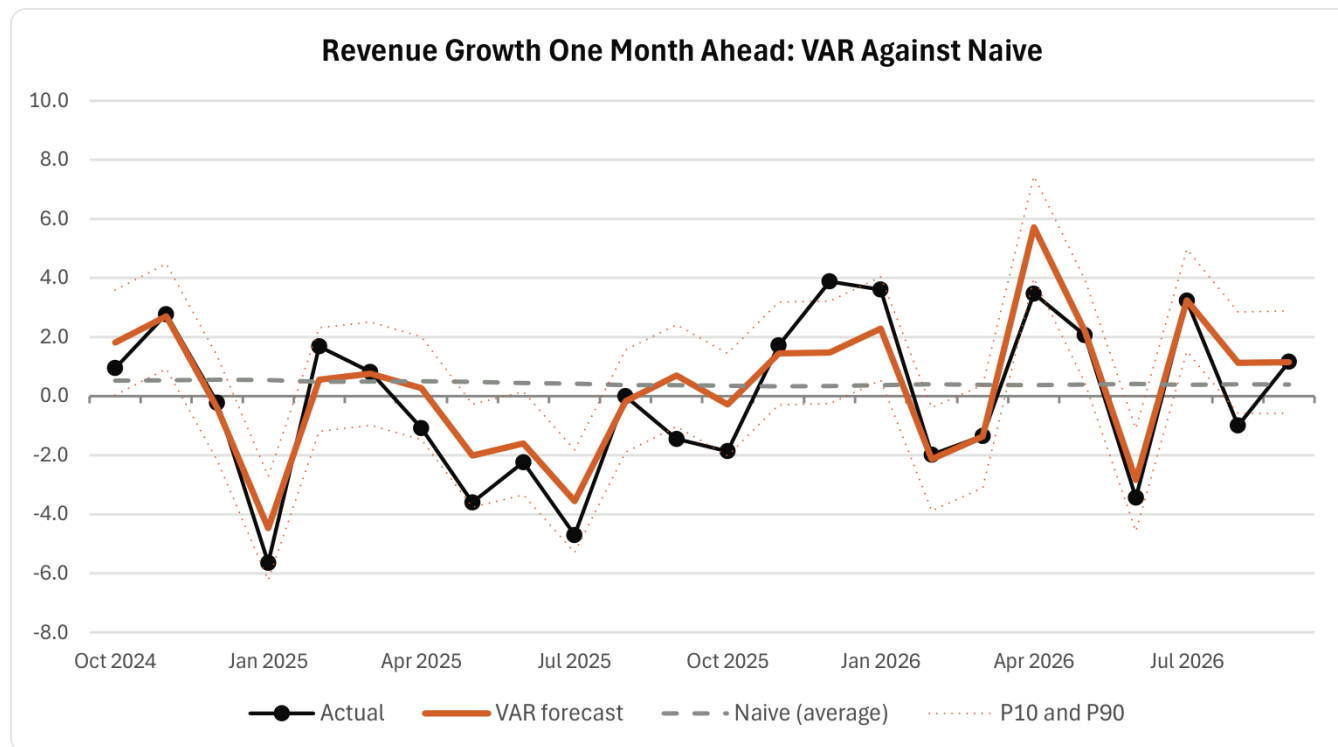
Revenue growth, October 2025 - September 2026	VAR	Naive (average)	Same as last month
RMSE (percentage points), all 12 months	2.46	2.52	3.34
RMSE, months 1-3	2.13	2.53	
RMSE, months 4-12	2.56	2.51	
Twelve months' revenue (\$ millions; actual 82.0)	80.0	79.2	

Over the whole year the VAR barely beats the naive average (2.46 against 2.52), and the reason is instructive. Its edge is in the first few months, when the order book already holds the answer (2.13 against 2.53). Beyond that, every model falls back to the average, and the VAR is no better than the average itself (2.56 against 2.51). Over the year it came in 2.0 million below actual revenue, against 2.8 million below for the naive trend. The bias was -0.31 points a month, and 10 of the 12 actual months fell inside the 80% band.

MAPE is the wrong score here. The backtest also reports MAPE, the average error as a percentage of the actual: 92%. That number is meaningless for growth rates. When the actual growth is near zero, any error is a huge percentage of it, and when it is exactly zero the percentage is infinite (XL Edge's MAPE leaves those months out, and says "n/a" if no others remain). Use RMSE or MAE for growth rates, and MAPE for levels.

One month ahead, 24 times

One holdout is one draw of luck. Sheet 7 *Backtest* also refits on everything known at the end of each month from September 2024 to August 2026, and forecasts the next month: 24 genuine one-month-ahead forecasts.



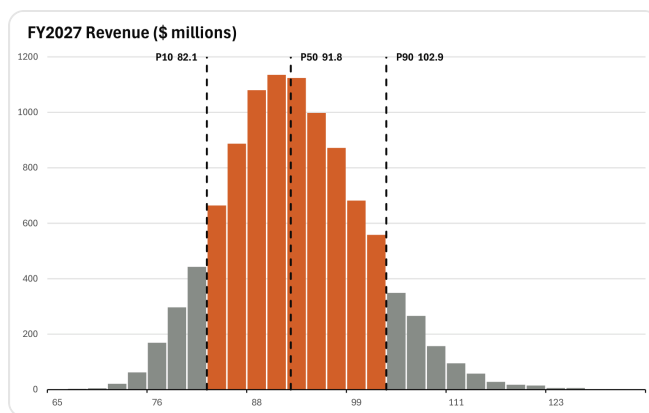
One month ahead, refitting every month: the VAR (orange) follows the swings the naive average (gray) cannot.

Here the VAR cuts the error by more than half: an RMSE of **1.19** points against 2.73 for the average and 3.60 for "same as last month". 20 of 24 actual months fall inside its 80% band (19.2 expected). This is the use a VAR is best at: the near term, where the leading indicators already contain the answer.

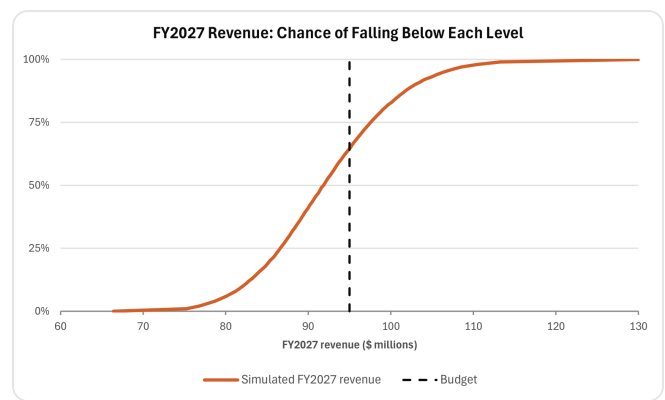
12 Step 8. From paths to the budget

Sheet 8 *Budget* turns each of the 10,000 simulated futures into FY2027 revenue: the sum of January to December 2027, months 4 to 15 of each path. The budget is the yellow cell `Budget`.

\$ millions	Value
FY2025 revenue, actual	81.9
FY2026 revenue, expected (actual to September, then the forecast)	84.5
FY2027 budget (growth of 12.4% over FY2026)	95.0
FY2027 on the point forecast	91.7
FY2027, simulated: mean	92.3
FY2027, simulated: P10 / P50 / P90	82.1 / 91.8 / 102.9
Chance of meeting the budget	35.5%
... the same, from the true model on the same random draws	35.6%



FY2027 revenue, 10,000 futures (*Outcome Histogram*). The P10 to P90 range is about 21 million wide.



The S-curve: at the budget line, read across to the chance of falling short, 64%.

Budget odds (sheet 8 Budget)

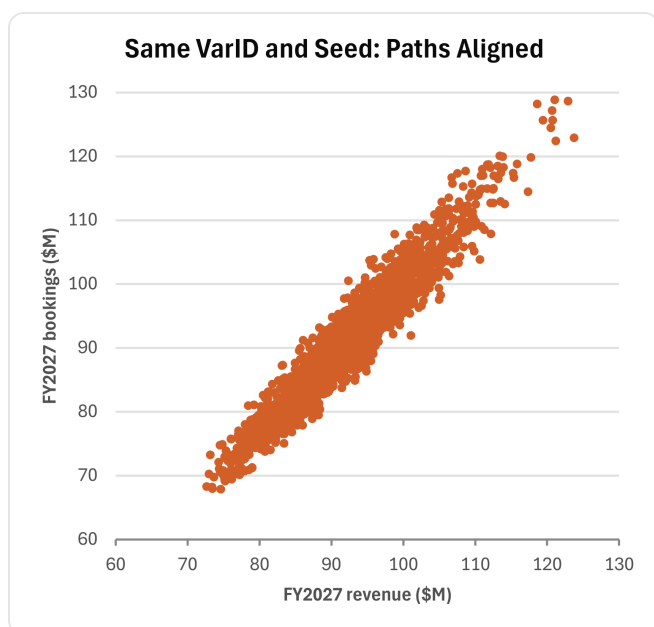
FY2027 budget (\$M)	85	88	90	92	95	98	100
Chance of meeting it	81%	69%	59%	49%	36%	23%	17%
True model, same draws	79%	67%	57%	49%	36%	25%	18%

WHAT TO TELL THE BUDGET COMMITTEE

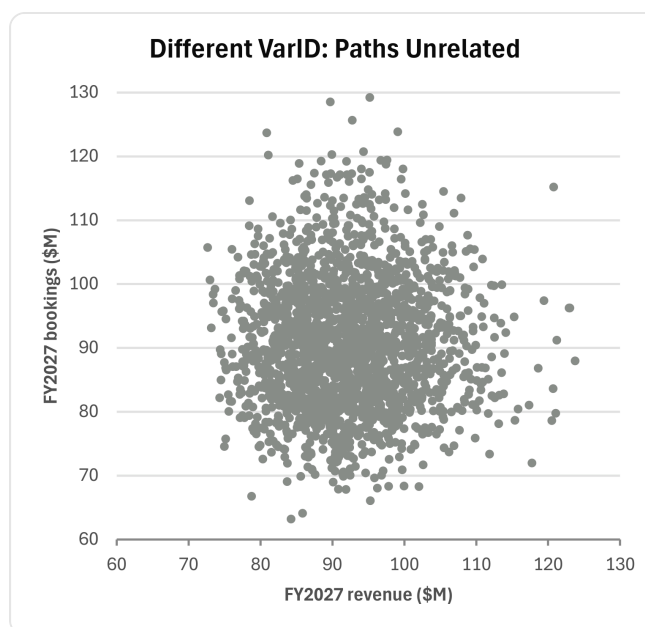
A 95-million budget is a stretch: about a one-in-three chance. The point forecast, 91.7 million, is close to a coin flip (49% at 92). A budget the company meets four years in five sits near 85 million (81%). None of these is wrong; they are different appetites for risk, and the table prices each one. The true model gives almost the same odds, so here the estimation error is small next to the uncertainty of the future itself.

Why the paths must share their random draws

Bookings were simulated twice: once with the revenue paths' VarID and Seed, once with a different VarID.



Same VarID and Seed: each dot is one future, bookings and revenue together.



Different VarID: bookings from one future paired with revenue from another.

FY2027	Same VarID and Seed	Different VarID
Correlation of bookings and revenue	0.96	-0.01
Book-to-bill, P10 to P90	0.94 to 1.03	0.82 to 1.18
Change in backlog (bookings – revenue), standard deviation (\$M)	3.1	13.0

Each series on its own is simulated correctly either way; the difference is only in how they combine. Pair bookings from one future with revenue from another, and anything built from both is wrong: the backlog's uncertainty comes out 4 times too large, and book-to-bill swings as no real order book does. Whenever a model uses more than one simulated series (a cash-flow model, a capacity plan), give every `fx.RiskVARSimλ` call the same VarID and Seed. The *Copula* guide makes the same point for correlated inputs.

13 Under the hood

Nothing in a VAR is a black box. Sheet 9 *Under the hood* rebuilds the functions' answers from first principles with ordinary Excel, and sheet *Checks* confirms each one to rounding.

What	How	Result
The fit	<code>fx.RiskVARDesignλ</code> lays out the regression (Y, then every lag and the constant); <code>LINEST(y, Z, FALSE)</code> on it gives the revenue equation, coefficient for coefficient.	identical
October's forecast	The revenue column of the Model block × the last two months of the three series and a 1, summed with <code>SUMPRODUCT</code> .	2.54%
October's P90	Forecast + <code>NORM.S.INV(0.9)</code> × $\sqrt{\Sigma(\text{revenue})}$: 2.54 + 1.282 × 1.346.	4.26%
The impulse response	Month 0: column 1 of <code>fx.RiskCholeskyλ(Σ)</code> . Month 1: A1 × that column, with <code>MMULT</code> . Month 2: A1 × month 1 + A2 × month 0, and so on.	0.065, 0.276
Stability	The companion matrix [A1 A2; I 0] (6 × 6) and <code>fx.RiskEigenvaluesλ</code> : the largest modulus.	0.769
Granger, bookings → revenue	Two regressions with <code>LINEST(..., FALSE, TRUE)</code> : with and without bookings' lags. $F = ((SSE_r - SSE_u) \div 2) \div (SSE_u \div 111)$ = ((672.2 – 201.1) ÷ 2) ÷ (201.1 ÷ 111).	130.05

One detail is worth knowing. The Granger F statistic is identical both ways, but its p-value is not: 8.1e–30 for one equation on its own against 1.8e–42 from `fx.RiskGrangerλ`. The VAR's test counts the residual degrees of freedom of all three equations, $K \times (T - m) = 333$ rather than 111, following statsmodels. With p-values this small it changes nothing, but near 5% the two can disagree.

A VAR in one line. Stacked, the whole model is $y(t) = c + A1 y(t-1) + A2 y(t-2) + e(t)$, with $e(t)$ drawn with covariance Σ . Every function in this guide is built from that line: the fit estimates c , $A1$, $A2$ and Σ ; the forecast runs it forward without e ; the simulation runs it forward with e ; the impulse response runs it forward from a single e .

14 Checking the model

Sheet *Checks* lists 20 tests, each shown as value, expected and difference, with a blank *Manual sign-off* cell for your ✓. Most are exact and differ only by rounding (about 10^{-12}). Three check the simulation against the formulas and pass when the difference is within three standard errors:

Simulation check	Simulated	Formula	Difference
October revenue growth: mean of 10,000 paths = the point forecast	2.554	2.538	1.6e–02
December 2027 revenue growth: P90 of the paths = the band's P90	4.244	4.213	3.1e–02
October: correlation of the bookings and revenue paths = the residual correlation	0.452	0.443	9.3e–03

The exact checks cover the LINEST rebuild, Σ , the October forecast and band, the impulse response at months 0 and 1, its 95% band (the same response as `fx.RiskVARIRFL`'s, with an upper limit 1.96 standard errors above it), the cumulative response, the stability root, the Granger F, the variance decomposition summing to 100%, the backtest RMSE, the last rolling forecast refitted directly, the true model's mean growth, and one simulated FY2027 rebuilt from its path.

15 Pitfalls and good practice

- **Make the series stationary first.** Model growth rates, or differences, not wandering levels. If levels are tied in the long run (bookings and revenue), consider a VECM for long horizons.
- **Keep the model small.** Each equation has $K \times p + 1$ coefficients: 3 series with 2 lags is 7, but 6 series with 4 lags is 25. With ten years of monthly data, three or four series and a few lags is typical.
- **Let the criteria propose and the residuals decide.** A model whose portmanteau test fails has missed a pattern; add a lag before trusting its dynamics.
- **The Cholesky order is an assumption.** Put the slowest-reacting or most "outside" series first (here the industry), and check whether your conclusions survive another order.
- **Read a response with its band.** `fx.RiskVARIRFBandsλ` shows which months the data can tell from zero. A narrow band does not prove the story behind it: the band assumes the order and the model.
- **Granger causality is about prediction,** and only among the series in the model.
- **Backtest against a simple benchmark,** from more than one origin. Expect the edge in the near term, and score growth rates with RMSE or MAE, not MAPE.
- **Turn growth into dollars path by path** (EXP of the running total), and sum the levels you need, as sheet *Paths (revenue)* does. Do not add up the growth band's percentiles.
- **Share the random draws.** Every `fx.RiskVARSimλ` call that feeds one model gets the same VarID and Seed.
- **Seasonal data:** pass `Season` (12 for monthly data) for centered seasonal dummies, or use seasonally adjusted series (`fx.RiskDecomposeλ`'s Adjusted column). **Outside drivers** (a price index, a policy rate) go in `Exog`, with their future values in the forecast's `FutureExog` (Part 2 shows it).
- **Using your own data:** paste monthly levels over *Data* D10:F130, oldest first, and rename the series in H9:J9. Everything recalculates. The *True model* comparisons then no longer apply.

16 Function reference

Data is $n \times K$: one column per series, oldest first, already stationary. Model is the block from `fx.RiskVARFit $\lambda$` (or typed by hand in its layout, or from `fx.RiskVECMToVAR $\lambda$`). Optional arguments are in brackets; skip one with an empty argument, e.g. `(Data, 2, , , , Names)`. All names start with `fx.Risk`.

Function	Output	Arguments and defaults
VARSelect	$(\text{MaxLags} + 3) \times 5$	(Data, [MaxLags]=10, [Type]="const", [Season], [Exog]). AIC, HQ, SC and FPE for 1 ... MaxLags on a common sample; the last row gives each criterion's choice.
VARFit	$(15 + m + K) \times (K + 1)$	(Data, [Lags]=1, [Type]="const" "trend" "both" "none", [Season], [Exog], [Names]). Least squares, one equation at a time; Σ with divisor $T - m$.
VARSummary	$K(m + 5) + 2(K + 1)$ rows	(Data, Model, [Exog]). Per equation: estimate, standard error, t, p; standard error, R^2 , adjusted R^2 . Then the residual covariance and correlation.
VARResids	$(T + 1) \times K$	(Data, Model, [Exog]). Names, then actual less fitted for rows $p + 1$... n .
VARStability	$(K p + 2) \times 4$	(Model). Verdict and largest modulus; then each root's real part, imaginary part and modulus, largest first. $K p$ up to 40.
VARPortmanteau	4×6	(Data, Model, [Lags]=10, [Alpha]=0.05, [Exog]). Box-Pierce and adjusted (Hosking) forms; $df = K^2(\text{Lags} - p)$, so Lags must exceed p .
Granger	5×6	(Data, Model, Cause, [Effect]=all others, [Alpha]=0.05, [Exog]). Cause and Effect by name or position, one or several. F and Wald χ^2 tests, and instantaneous causality.
VARIRFA	$(H + 3) \times (1 + K^2)$	(Model, [Horizon]=10, [Impulse], [Orthogonal]=TRUE, [Cumulative]=FALSE, [Order]). Columns "shock $\rightarrow$ response", periods 0 ... H ; $K + 1$ columns with one Impulse.
VARIRFBands	$(H + 3) \times 5$	(Data, Model, Impulse, Response, [Horizon]=10, [Level]=0.95, [Orthogonal]=TRUE, [Cumulative]=FALSE, [Order], [Exog]). One shock $\rightarrow$ response pair: Period, Response, Lower, Upper, Std error. Asymptotic (delta-method) standard errors, as statsmodels reports them.
VARFEVDA	$(K H + 2) \times (K + 2)$	(Model, [Horizon]=10, [Response], [Order]). Shares of forecast-error variance, horizons 1 ... H ; $(H + 2) \times (K + 1)$ with one Response. Each row sums to 1.
VARForecast	$(H + 1) \times (1 + 3K)$	(Data, Model, Horizon, [Variable], [Lower]=0.1, [Upper]=0.9, [FutureExog], [Labels]). Point forecast and band for each series; with Variable, Period / Forecast / P10 / P90.
VARSim	Trials $\times$ Periods	(Data, Model, Periods, [Variable]=1, [DegFree], [Trials]=10000, [VarID]=1, [Seed]=0, [FutureExog]). One series' paths, drawing all K shocks; same VarID and Seed align the series. DegFree > 2 for Student t shocks.
VARBacktest	$(\text{Holdout} + 6) \times 6$	(Data, Holdout, [Lags]=1, [Type], [Variable]=1, [Lower], [Upper], [Season], [Names], [Exog]). MAPE (over the non-zero actuals), RMSE, MAE, Bias, Coverage; then Period, Actual, Forecast, Error and the band. One forecast origin; with Exog, its held-out rows are the future values.

Helpers used in this guide

Function	Output	Arguments
<code>VARDesignλ</code>	$(1 + n - p) \times (K + m)$	(Data, Lags, Type, Season, Exog, Skip, Names), all required ("" for no Exog or Names, 0 for no Season or Skip). The regression: Y, then the lags and deterministic terms.
<code>Choleskyλ</code>	$K \times K$	(Matrix). Lower-triangular factor of a covariance matrix.
<code>Eigenvaluesλ</code>	$(n + 1) \times 3$	(Matrix). Real, imaginary and modulus of any real square matrix, largest first.
<code>UnitRootλ</code>	5×9	(X, [Trend]="c" "ct", [Lags], [Alpha], [Name]). ADF, Phillips-Perron and KPSS.
<code>ChartFanλ</code>	$(\text{Periods} + 2 \text{ or } 3) \times 7$	(Paths, [Lower], [Upper], [Name], [Start], [Plan], [Labels]). Fan-chart data for <i>Fan Chart (Time Series)</i> .

References: Sims (1980), "Macroeconomics and Reality", *Econometrica* 48(1); Granger (1969), *Econometrica* 37(3); Lütkepohl (2005), *New Introduction to Multiple Time Series Analysis*, ch. 2-4; Hosking (1980), *JASA* 75; Pfaff (2008), "VAR, SVAR and SVEC Models: Implementation Within R Package vars", *Journal of Statistical Software* 27(4); Zivot and Wang, *Modeling Financial Time Series with S-PLUS*, ch. 11; Woodward, Sadler and Robertson (2022), *Time Series for Data Science*, ch. 10. The VAR functions are validated against statsmodels. The example's history is simulated and its figures are illustrative, not any real company's. XL Edge and the fx.Risk... functions: © 2026 W Phillips, edgewisedata.com, MIT License.