Which Macro and Market Series Predict Which? A Point-in-Time Dependency Graph

Will Sackfield, Open Family Office · https://open-familyoffice.com/research/dependency-graph

Abstract

Practitioners often draw maps of which economic and market series drive which, but such maps rarely state how many links were searched, when each input became known, or whether a link holds out of sample. We build such a map for 33 public series: 14 macroeconomic series from FRED and 19 market prices. We test every pair with a Granger regression at the slower of the pair's two frequencies, with errors that allow for autocorrelation. We adjust the 5,090 Granger tests jointly to a 10% false discovery rate, and keep an edge only if its sign holds in three separate periods and in a held-out final fifth of the data. Of 1,056 ordered pairs, 162 keep a usable edge. The largest group, 61 pairs, is market prices that predict macroeconomic series, and the broad dollar index is predicted by 18 other members. Two of our own earlier builds show how easily timing creates edges. Dating each value by its release made a series that FRED publishes a day late look as though the day's market predicted it, at adjusted p-values as small as 10^{-99}. Dating each value by the day it describes removed those edges. A second method, PCMCI, which conditions on the other series, finds 66 usable pairs on the dates its members share. Only 22 of them are among the Granger pairs, although all 22 agree in sign. So most of the Granger graph is not kept when the other series are held constant.

Status: Published 2026-11-28. Changes since are in the addenda at the end.

1 Introduction

A map of which series help to predict which is attractive to an investor. It promises to say which inputs a strategy should watch, and which apparent signals are only echoes of a common driver. Such maps are common in practice. One recent example trains a panel of forecasting models on 33 macro and market series, and draws a graph from the models' feature importances (Sackfield, 2026). Maps of this kind seldom report three things that decide whether their links are real. They seldom say how many candidate links were searched, whether each input was used only after it became public, or whether a link survives in data that the search did not see.

Each of the three has a known remedy. A search over many links needs a correction for multiple testing, a problem that affects many questions in finance, such as the selection of funds and factors (Harvey & Liu, 2020). A false discovery rate correction keeps the expected share of false links among the reported ones below a chosen level (Benjamini & Hochberg, 1995). Macroeconomic data are revised and published with a lag, so a test must read each value as it was first known. A link that is found in one period should be

checked in others, and in a period held out of the search. Methods that condition on many series at once can also remove links that a pairwise measure reports between autocorrelated series (Runge et al., 2019).

This paper applies those remedies to the same 33 series, and reports what survives. Our contributions are:

- A dependency graph of 33 public series, built point-in-time, with each of its 9,150 tests counted. Each method's tests are adjusted jointly for a 10% false discovery rate, and each edge is checked for a stable sign in three periods and out of sample.
- A description of the 162 ordered pairs that keep a usable edge: which kinds of series predict which, at what frequency and lag, and which series are hubs.
- A measured example of how the dating rule for each value creates or removes edges. Three builds of the same graph, which differ only in that rule, give 232, 181 and 162 usable pairs.
- A comparison with PCMCI, which conditions on the other series: it keeps 22 of the 162 Granger pairs and finds 44 others, and the shared pairs agree in sign.

The work bears on the project's first research question, how to evaluate strategies so that few false discoveries pass (`research-questions.md`, question 1), and on question 8, which this paper adds.

2 Related work

Two lines of work meet in this graph: tests of predictive links between time series, and corrections for searching many of them. The first test of predictive links is the regression named after Granger. A series X is said to Granger-cause Y when the past of X improves the forecast of Y beyond the past of Y alone. The word "cause" is a misnomer that the field keeps, because the test measures only an improvement in prediction. Runge et al. (2019) show two failures of the simpler approaches. Bivariate measures, such as a correlation between two series, report many false links when the series are autocorrelated. A full conditional test, which conditions on every other series, avoids them but loses power as series are added. Their method, PCMCI, first selects a small set of candidate parents for each series. It then tests every pair at every lag, and conditions each test on the parents that its first step selected. In their simulations it controls false positives near the nominal level and keeps more power than the full conditional test. Our Granger test conditions on the target's own past but on no third series, so it is closer to the bivariate approaches, and the comparison with PCMCI checks it. They do not correct for the number of links tested, although their software offers a Benjamini–Hochberg option.

The second line is about the search itself. Benjamini and Hochberg (1995) define the false discovery rate as the expected share of false rejections among all rejections, and give a step-up procedure that controls it when the test statistics are independent. Harvey and Liu (2020) compare such corrections, including the Benjamini–Yekutieli form that holds under any dependence, on equity strategies, published anomalies and mutual funds. They show that the choice changes both the false discoveries and the true ones that are missed. Our graph is a search of the same kind, and we use their lesson: the threshold must account for every test that was run, including the ones that failed.

The motivating example is a post by the project's maintainer (Sackfield, 2026). It ranks 33 series by their centrality in a graph of feature importances, with crude oil, the dollar

index, the VIX and the slope of the yield curve at the top. It groups the series as 15 macroeconomic and 18 price series. We count the 13-week bill yield, a market quote, with the market series. Its graph is drawn from importances, which are not tests, and it reports neither a correction nor an out-of-sample check. Our graph keeps its members and replaces its method. The regressions use the statsmodels library (Seabold & Perktold, 2010), with errors in the form of Newey and West (1986), which allow for heteroskedasticity and autocorrelation. Finally, a held-out check of the sign is weaker than the out-of-sample comparison of forecasts by which Welch and Goyal (2008) found that popular predictors of the equity premium fail, and we return to this in the Limitations.

3 Method

Members. The graph has 33 members, the panel of the motivating post (Sackfield, 2026). Fourteen are FRED series. Seven are rates, spreads or levels quoted daily: the ICE BofA US high-yield option-adjusted spread, WTI crude oil, the ten-year Treasury yield, the broad dollar index, the ten-year minus two-year spread, the ten-year breakeven inflation rate and the VIX. The others are consumer prices, the federal funds rate, GDP, industrial production, M2, vehicle miles travelled and unemployment. Nineteen are daily market prices from Yahoo Finance: equity, bond, property and crypto funds, two crypto coins, two exchange rates, four commodity futures and the 13-week bill yield. Twenty-six members are daily, six monthly and one quarterly. Appendix B lists them with their identifiers.

Changes. Each member is made stationary by a log change for prices and levels that stay above zero, and by a plain difference for rates, spreads and the two oil prices, whose stored values include a day below zero.

Point-in-time values. A value is point-in-time when a test uses it only as it was known at the time. ALFRED, FRED's archive, keeps past versions of each observation, called vintages, each with a real-time start date. We read each observation in its earliest stored vintage, so that a revision after that date does not enter the test. ALFRED's vintages start only when a series was added to the archive: in 1972 for consumer prices and in 1991 for GDP, by the review of FRED's API. Observations from before that date are therefore read in a later, revised vintage. Each observation is then placed at the end of the day, month or quarter that it describes. A monthly series such as consumer prices, dated by FRED on the first day of its month, is therefore read at that month's end. An observation enters the build only if it was published by the build's time. Market prices are read on one share basis, so that a split does not appear as a return. The section on dating rules explains why we date values by the period they describe rather than by their release.

Pairs and frequencies. Each pair is tested at the slower of its two members' frequencies: on weekdays, at month ends or at quarter ends. A monthly series read every day would repeat one value for a month, which inflates its persistence and can create links. A pair uses only the dates on which both members have a change.

The test. For a source X and a target Y at lag L, the test fits

$$Y_t = c + \sum_{k=1}^{L} a_k Y_{t-k} + \sum_{k=1}^{L} b_k X_{t-k} + \varepsilon_t,$$

where c is a constant, a_k and b_k are coefficients and ε_t is the error. The p-value is that of a Wald test that every b_k is zero, with the Newey and West (1986) covariance. We set its

number of lags equal to the test's lag L. That choice is ours, and the Newey–West paper gives little guidance on it. The edge's size is $\sum_k b_k$, and its sign is the sign of that sum. Each ordered pair is tested at each lag L from 1 to 5, so one pair in both directions gives ten tests.

Out of sample, and stability. The last fifth of each pair's dates is held out of the fit. An edge is *held* when the sign of the size, fitted on the held-out dates alone, matches the sign fitted on the rest. The rest is cut into three equal periods. An edge is *stable* when its sign is the same in each of them.

The correction. The tests of each method, across the whole graph, enter one Benjamini–Hochberg adjustment (Benjamini & Hochberg, 1995). Its output for each test is an adjusted p-value, written q: the smallest false discovery rate at which the test would count as a discovery. An edge is a *discovery* when q is at most 0.10. An edge is *usable* when it is a discovery, stable and held. A pair is usable when at least one of its five lags gives a usable edge. Every test is also recorded as a trial in the project's trial ledger, so the size of the search is on record.

The second method. PCMCI (Runge et al., 2019) runs once per frequency over the members read at that frequency, so each link is conditioned on the others. It uses the partial-correlation test, lags of up to 5, and a level of 0.05 for keeping a parent in its first step. A joint run uses only the dates that all its members share, and leaves out any member with less than ten years of changes. It gives one edge per ordered pair and lag, with the same stability and hold-out checks, and its own adjustment.

Builds. The graph was built four times on 2026-11-28, from the same stored data. The first three differ in the rule that dates each value, and are the subject of the section on dating rules. The fourth, as of 22:32:41 UTC, changes only PCMCI's sample, to the dates its members share. Its Granger results are identical to the third build's, and the other sections report the fourth. Its 9,150 tests are 5,090 Granger tests on 509 of the 528 pairs, and 4,060 PCMCI tests on 406 pairs. The rest were skipped, because a pair had too few common dates for five lags in each period, or because a member had too short a history for a joint run.

Figures. Figure 2 is animated: it steps through the first three builds, because its point is the change from one dating rule to the next. Figures 1, 3 and 4 show the fourth build and do not move.

4 Which pairs keep a usable edge?

The question is how much of the full search survives the correction and both checks. Of the 5,090 Granger tests of the final build, 886 are discoveries at a 10% false discovery rate. Of these, 603 keep their sign in all three periods and 651 keep it out of sample. The 471 edges that pass both checks belong to 162 ordered pairs, out of 1,056 possible (Figure 1). The checks matter: about half of the discoveries fail at least one of them, which suggests that a corrected p-value alone overstates how many links would hold in new data.

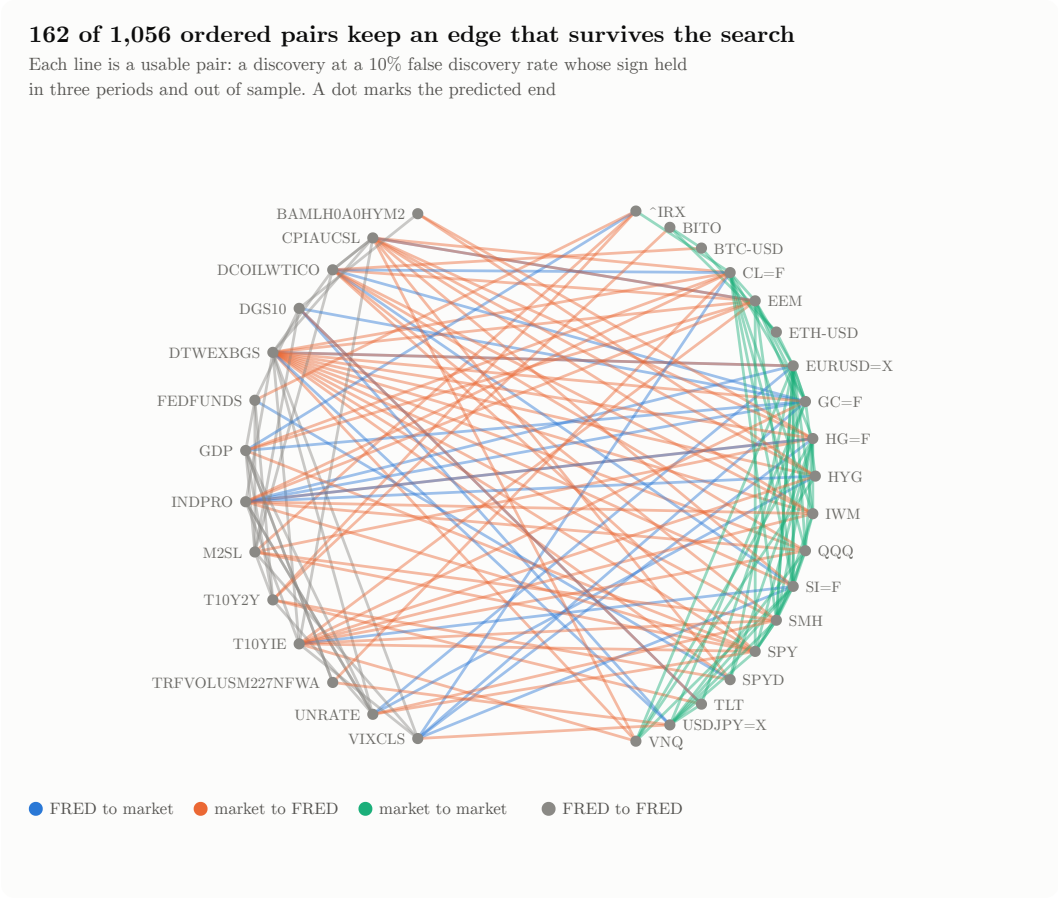

Figure 1. The usable Granger graph of the final build. The 33 members sit on a circle, the 14 FRED series on the left and the 19 market series on the right. Each line is an ordered pair with at least one usable edge, and a dot marks its predicted end. Colour shows the datasets at the two ends.

The usable pairs are not spread evenly over the kinds of pair (Table 1). Market prices that predict FRED series form the largest group, with 61 of the 162 pairs, and market prices that predict other market prices form the next, with 50. Only 23 pairs run from a FRED series to a market price. That imbalance is what a reader would expect if markets anticipate the macroeconomic series, which are measured with a delay. It is also what timing differences within a day would produce, as the Discussion explains.

Table 1. Usable ordered pairs in the final build, by the datasets at their two ends.

Kind of pair	Usable pairs
Market to FRED	61
Market to market	50
FRED to FRED	28
FRED to market	23
All	162

Most usable pairs are daily, and most of their strongest edges are at the shortest lag (Table 2). Of the 162 pairs, 103 are tested daily, 48 monthly and 11 quarterly. The strongest edge of 70 pairs is at a lag of one period, and the sign is positive for 101 pairs

and negative for 61. Short lags dominate because daily data give the most dates and therefore the most power.

Table 2. Granger tests and usable results in the final build, by the frequency at which the pair is tested.

Frequency	Tests	Usable edges	Usable pairs
Daily	3,250	333	103
Monthly	1,590	116	48
Quarterly	250	22	11

The strongest pairs are concentrated on one member (Table 3). Nine of the twelve strongest pairs predict the broad dollar index or are predicted by it. The equity and property funds each predict a fall in the dollar index one or two days later, the VIX predicts a rise, and the dollar index predicts a fall in the euro's dollar price. The monthly pairs that reach the top are a commodity and an inflation series: the crude oil future predicts consumer prices one month later, and consumer prices predict crude oil three months later with a negative sign.

Table 3. The twelve strongest usable pairs of the final build, with the lag and sign of each pair's strongest edge. Every adjusted p-value is below 10^{-7}.

Source	Target	Frequency	Lag	Sign
EEM	DTWEXBGS	daily	2	−
DTWEXBGS	EURUSD=X	daily	5	−
IWM	DTWEXBGS	daily	1	−
SPY	DTWEXBGS	daily	1	−
SMH	DTWEXBGS	daily	1	−
VNQ	DTWEXBGS	daily	1	−
QQQ	DTWEXBGS	daily	1	−
BTC-USD	BITO	daily	3	+
VIXCLS	DTWEXBGS	daily	1	+
CL=F	CPIAUCSL	monthly	1	+
CPIAUCSL	DCOILWTICO	monthly	3	−
SPYD	DTWEXBGS	daily	2	−

The hubs of the graph follow from these pairs (Figure 3). The dollar index is predicted by 18 members and predicts 3. The euro, silver and consumer prices are each predicted by 11. The equity funds are the main predictors: the semiconductor fund, the emerging-market fund and the S&P 500 fund each predict 11 members, and the small-company fund predicts 10 and is predicted by none. This ranking differs from the motivating post's, in which crude oil ranked first and the dollar index second (Sackfield, 2026). Both put the dollar index near the top, although the post measured importance and this graph counts tested links.

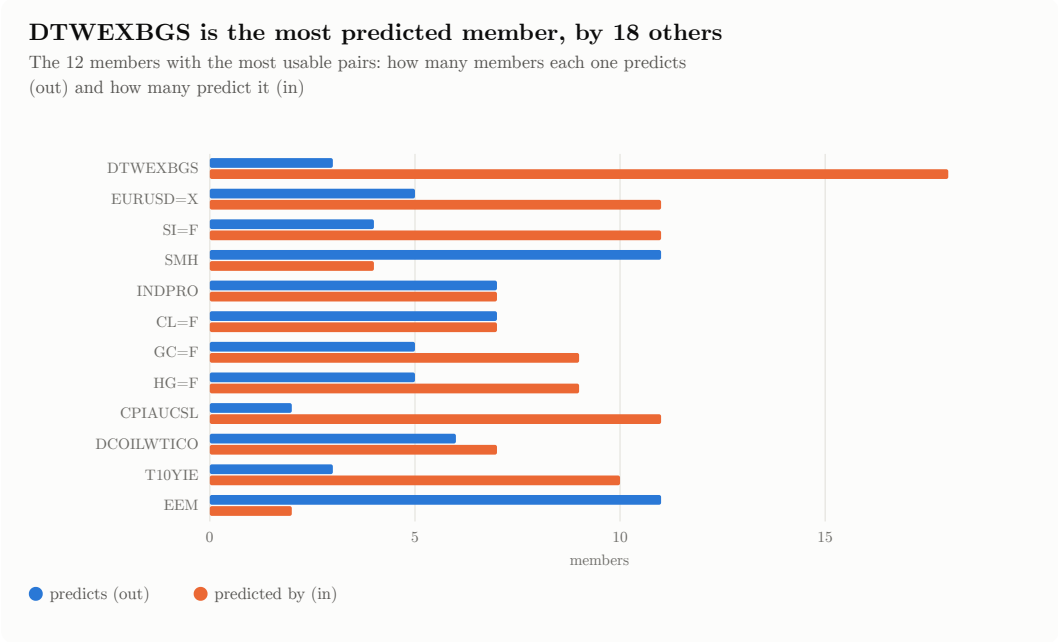

Figure 3. The 12 members with the most usable pairs in the final build: the number of members each one predicts, and the number that predict it.

5 How does the dating rule create edges?

The question is whether the rule that dates each value can, by itself, create strong edges. We built the same graph three times from the same data, and changed only that rule. In the first build, each value counted from its release: a FRED vintage from the day after its publication date, and a back-filled market close from the day after its own date. In the second, a market close counted from the end of its own day. In the third, every value is placed on the day it describes, in its first vintage.

The first build gave 232 usable pairs, and its strongest links were artefacts of timing (Figure 2, Table 4). In the stored vintages, a day's value of the VIX or of the high-yield spread has that same day as its real-time start, so it counted a day before the same day's market close. The VIX therefore appeared to predict the Nasdaq-100 fund the next day, with an adjusted p-value near 10^{-76}, although both moved on the same day. The second build moved each close to its own day, which removed those links, with adjusted p-values of 0.62 and 0.54. It created others in the opposite direction. A day's ten-year yield has the next day as its real-time start, so the Treasury bond fund appeared to predict it, with an adjusted p-value near 10^{-80}. Every such edge in the first two builds measures a publication lag, not a relation between the series.

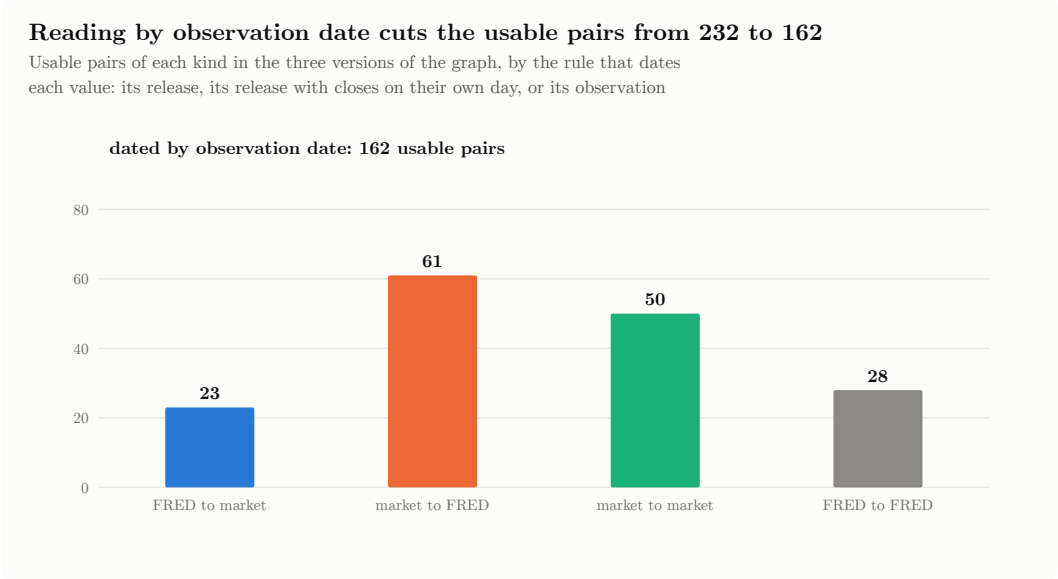

Figure 2. The usable pairs of each kind in the three builds. The animation steps through the builds in order, and rests on the third. The second build differs from the first only for market closes, and the third dates every value by the day it describes.

Table 4. The adjusted p-value q of four lag-1 Granger edges in each build. A small q in one build and a large one in the third marks an edge created by the dating rule.

Edge	By release	Close on its own day	By observation
VIXCLS to QQQ	1.5×10^{-76}	0.62	0.89
BAMLH0A0HYM2 to IWM	9.5×10^{-69}	0.54	0.55
TLT to DGS10	4.5×10^{-99}	6.4×10^{-80}	0.11
CL=F to DCOILWTICO	3.6×10^{-7}	0.69	0.62

The third build removed all four. Dating by observation places two series that describe the same day on the same date, whatever their publication lags, so a lag in publication can no longer appear as a lag in the data. The cost is that the graph no longer says what was known before what. An edge in it states that one series' value for a period predicts another's value for a later period, which is the question an economic map asks. A strategy that trades on an edge must still wait for both values to be published. The count of usable pairs fell from 232 to 162 across the three builds, and the kinds of pair shifted. Pairs from FRED to market fell from 60 to 23, which suggests that many of them were timing artefacts of the kind in Table 4.

6 Does PCMCI agree with the Granger graph?

The question is whether a method that conditions on many series at once keeps the same links. Its first attempt could not answer. In the third build, PCMCI ran over the full calendar from the earliest member's first value, in 1919, so the first two of its three periods held no market data, and no edge was stable. The fourth build runs each joint fit on the dates that its members share, and leaves out the three members with less than ten years of changes: the high-yield spread, the Bitcoin futures fund and Ether. Its earliest fit

then starts on 23 October 2015, and its fits end by 9 September 2024, before the held-out fifth.

On that sample, PCMCI keeps 143 discoveries among its 4,060 tests, of which 96 are stable and 97 are held, and 67 usable edges belong to 66 ordered pairs. The two methods share only 22 usable pairs (Figure 4). Granger finds 140 pairs that PCMCI does not, and PCMCI finds 44 that Granger does not. Where both find a pair, they agree: all 22 shared pairs have the same sign, and 16 the same lag.

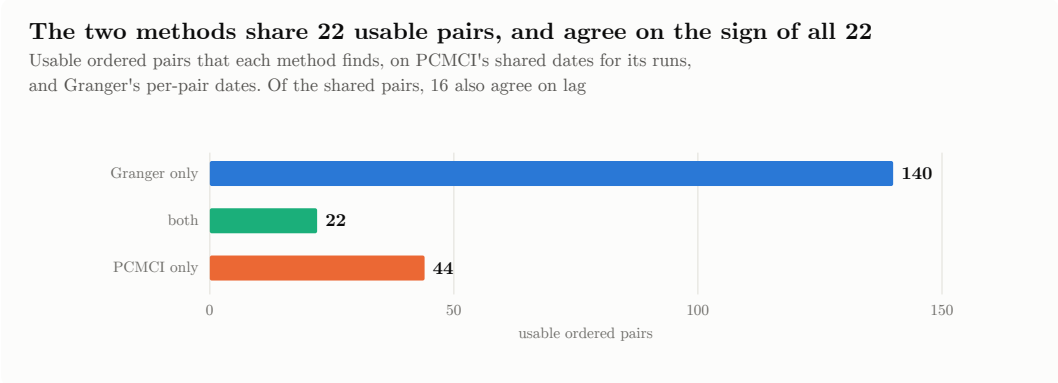

Figure 4. The usable ordered pairs that each method finds in the final build, alone and together.

The shared pairs are the most plausible part of the graph. They include crude oil, both the futures price and FRED's spot price, predicting consumer prices. They also include the bill yield predicting the federal funds rate, and the bond fund predicting the ten-year yield. Further shared pairs are the dollar index and the euro, in both directions, and several equity funds predicting copper. Most of the 140 Granger-only pairs, such as most of the one-day leads into the dollar index, are not kept once the other series are held constant. That pattern fits the finding of Runge et al. (2019) that bivariate measures report false links among autocorrelated series. Our reading, that a common driver explains many of these pairs, is our own. The comparison has a caveat in sample: PCMCI's joint runs fit on about nine years of shared dates, whereas many Granger pairs fit on decades. So some Granger-only pairs may reflect the longer sample rather than a common driver.

7 Discussion

Four results stand out. The search is large, and the corrections remove most of it: 5,090 Granger tests give 886 discoveries and 162 usable pairs. The surviving graph is led by markets that predict macroeconomic series and by one hub, the dollar index. A method that conditions on the other series keeps only 22 of those pairs, although it agrees with all of them in sign. The dating rule can create edges as strong as any real one, which suggests that a map of this kind is only as trustworthy as its treatment of time.

The direction from markets to macroeconomic series has two explanations, and this graph cannot separate them. The first is economic: prices aggregate information quickly, and a macroeconomic series measures the same conditions later. The second is mechanical, and it is our own explanation of the dollar index as a hub. The Federal Reserve's notes on its H.10 release describe its exchange rates as noon buying rates in New York, and the broad

dollar index is published with that release. If the index uses rates taken at noon, an afternoon move in equities is in that day's close but in the next day's index, which makes the equity fund appear to lead the index by one day. The same mechanism can make Bitcoin, which trades every day, appear to lead a Bitcoin futures fund that trades only on weekdays. The many one-day lags into the dollar index, almost all with a negative sign, fit this second explanation. A strategy should not rest on such an edge until it is tested on values sampled at the same moment.

The monthly links are less exposed to this problem. Crude oil predicts consumer prices a month later, which is a plausible pass-through of energy prices into the price index. The negative link from consumer prices to crude oil three months later has no such reading, and may be a false discovery. The adjustment controls the expected share of false discoveries, not their number, and at a 10% rate several of the 162 pairs are expected to be false.

These results support three practices for any graph of predictive links. Every test should be counted and adjusted together, because a graph is a search. A pairwise link should be checked by a method that conditions on the other series before a strategy rests on it. Each value should be dated by the period it describes and read in its first vintage, and the dates within a day should be checked, because timing alone can create edges with adjusted p-values near 10^{-99}.

8 Limitations

- **Dependent tests.** The procedure of Benjamini and Hochberg (1995) is proved to control the false discovery rate for independent test statistics. This graph's tests share series, lags and dates, so they are strongly dependent, and the 10% rate is not guaranteed. The Benjamini–Yekutieli form would be stricter. A bootstrap method such as StepM, which estimates the dependence between tests (Romano & Wolf, 2005), would control a stricter error rate.
- **Prediction, not cause.** A Granger edge states that one series improves the forecast of another. It does not exclude a third series that drives both, and the graph tests no structural model.
- **Sampling within the day.** Members are sampled at different times of day, from a noon rate to an afternoon close to a 24-hour crypto day. Edges at one day's lag between such members may reflect that difference.
- **Short or recent members.** The stored history of the high-yield spread covers only a few years, and the Bitcoin futures fund starts in 2021, so their pairs are tested on few dates or skipped.
- **Vintages start late.** ALFRED's vintages start when a series entered the archive, in 1991 for GDP. Earlier observations are read in a revised vintage, so the early part of a pair's sample is not fully point-in-time.
- **Different samples for the two methods.** PCMCI's joint runs need dates that all their members share, so they fit on about nine years and leave out three young members. The Granger pairs use each pair's own, often longer, dates. Some disagreement between the methods may come from the samples.
- **Linear tests, and a weak hold-out check.** The regressions detect only linear predictive links at lags of up to five periods. The hold-out check asks only whether the sign repeats. It does not ask whether the edge improves a forecast out of sample

against a simple benchmark, which is the stricter test that Welch and Goyal (2008) apply.

- **Links within a period.** Dating by observation puts same-day movements on the same date, where a test of lags does not see them. Links within one period need a different method.

9 Conclusion

Of 1,056 ordered pairs among 33 public macro and market series, 162 keep a predictive edge after a joint false discovery rate correction and checks of stability and of a held-out period. The surviving graph is led by market prices that predict macroeconomic series, with the dollar index as its main hub. Several of its strongest edges may still reflect the times of day at which the series are sampled. The rule that dates each value proved decisive: dating by release created edges with adjusted p-values as small as 10^{-99}, and dating by the period described removed them. A method that conditions on the other series keeps 22 of the 162 pairs, all with the same sign. That smaller set is the part of the graph on which a strategy can most safely rest.

References

Benjamini, Y., & Hochberg, Y. (1995). *Controlling the false discovery rate: A practical and powerful approach to multiple testing.* Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300. doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Harvey, C. R., & Liu, Y. (2020). *False (and missed) discoveries in financial economics.* Journal of Finance, 75(5). doi.org/10.1111/jofi.12951.

Newey, W. K., & West, K. D. (1986). *A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix.* NBER Technical Working Paper 55, published in Econometrica, 55(3), 1987. nber.org/papers/t0055.

Romano, J. P., & Wolf, M. (2005). *Stepwise multiple testing as formalized data snooping.* Econometrica, 73(4), 1237–1282. doi.org/10.1111/j.1468-0262.2005.00615.x.

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., & Sejdinovic, D. (2019). *Detecting and quantifying causal associations in large nonlinear time series datasets.* Science Advances, 5(11), eaau4996. doi.org/10.1126/sciadv.aau4996.

Sackfield, W. (2026). *Economic causality: Finding out what affects what in the markets.* Subsack. sackfield.substack.com/p/economic-causality.

Seabold, S., & Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with Python.* Proceedings of the 9th Python in Science Conference, 92–96. doi.org/10.25080/Majora-92bf1922-011.

Welch, I., & Goyal, A. (2008). *A comprehensive look at the empirical performance of equity premium prediction.* Review of Financial Studies, 21(4), 1455–1508. doi.org/10.1093/rfs/hhm014.

Appendices

Appendix A: Reproduce

The graph's data are private: the series stay on the project's Modal Volume, and only statistics of the edges are published. With access to the Volume, the graph is built and summarised with:

```
modal run scripts/modal_app.py::build_graph --name macro
modal volume get open-familyoffice-data graphs/macro.jsonl <folder>/
uv run python scripts/dependency_graph_results.py <folder>/macro.jsonl
uv run python scripts/paper_figures.py
```

The build has no random seed. The graph's code is `open_familyoffice/graphs.py` and `open_familyoffice/graph.py`. The four builds ran at the commits of #542, #552, #556 and #579, the last with PCMCI's shared window from #573. Each build's tests are in the trial ledger under the family `graph-macro`: 6,090, 6,090, 6,800 and 9,150 trials. The number of skipped pairs in each build comes from the build's own output, which the stored graph does not keep: 447, 447, 376 and 141. So do the three members left out of the fourth build's joint PCMCI runs.

Appendix B: Members

Member	Dataset	Frequency	Change
BAMLH0A0HYM2	FRED	daily	difference
CPIAUCSL	FRED	monthly	log
DCOILWTICO	FRED	daily	difference
DGS10	FRED	daily	difference
DTWEXBGS	FRED	daily	log
FEDFUNDS	FRED	monthly	difference
GDP	FRED	quarterly	log
INDPRO	FRED	monthly	log
M2SL	FRED	monthly	log
T10Y2Y	FRED	daily	difference
T10YIE	FRED	daily	difference
TRFVOLUSM227NFWA	FRED	monthly	log
UNRATE	FRED	monthly	difference
VIXCLS	FRED	daily	log
^IRX	Yahoo Finance	daily	difference
CL=F	Yahoo Finance	daily	difference
BITO, BTC-USD, EEM, ETH-USD, EURUSD=X, GC=F, HG=F, HYG, IWM, QQQ, SI=F, SMH, SPY, SPYD, TLT, USDJPY=X, VNQ	Yahoo Finance	daily	log

Appendix C: Who did the work

The maintainer chose the members, asked for the graph and chose to date values by observation. Claude Opus 5.5, as an agent of the project, wrote the graph's code, ran the three builds, found the timing artefacts and the flaw in PCMCI's sample, and wrote this paper and its figures.