The Backtest Library #1

What Peter Lynch's PEG Rule Actually Does in Real Markets

One of the few places where Lynch comes close to offering an actual number is valuation. He is suspicious of buying a wonderful company at any price, and equally suspicious of buying a cheap company that is quietly dying. His answer is the relationship between the price you pay and the growth you get.

This library starts with Peter Lynch for one reason: the premise of this site and the premise of his book are the same premise - that an ordinary person, without an office on Wall Street, can beat the market by doing a few simple things carefully.

This report is not written for fund managers. It exists to answer a plain question: if you take a stock-picking method that has been circulating in books for decades and actually follow it, step by step, does it make money?

The short answer, so you do not have to wait for it: no. And it fails in an interesting way. It does not blow up your account. It grinds for thirteen years and lands roughly where a basket of stocks picked at random would have landed, a few points a year short of the index and a long way short of the growth stories it was meant to catch.

Each section below is built the same way. First what we did, then one short passage headed "What it means for you", in plain language. If you only read those passages, you will still get the argument.

The PEG idea

Lynch's most-quoted rule of thumb fits in one sentence: take a company's price-to-earnings ratio and divide it by the rate at which its earnings are growing. A company trading at 15 times earnings and growing at 15% a year has a PEG of 1.0. His rough reading of that number is simple. Around 1 is fair. Below 1, you may be getting growth for less than it is worth. Well above 1, you are paying a premium for a story that has to be justified by something beyond the current numbers.

In one line: the P/E is what the growth costs, the growth rate is what backs it up, and PEG is the price per unit of growth. Lower means you are paying less for the same growth.

There is a precondition that gets skipped whenever the rule is repeated. The ratio only means anything for one kind of company, and that is not our judgment: it is the reason Lynch split stocks

into six groups. Fast growers, cyclicals, turnarounds, slow growers, asset plays, and the steady growers. Each group's P/E is read differently, and PEG holds for exactly one of them.

Start with the one where it holds. A fast grower is a company whose earnings climb a steady 15% to 25% a year by opening new stores or taking share. Lynch's example is Body Shop, growing earnings 20% to 30% a year. His test for this group is blunt: "one test of whether a stock is overpriced is that the P/E should not exceed the earnings growth rate." He also put a ceiling on the growth rate itself: "even the fastest-growing companies rarely manage 25% earnings growth, and 40% is extremely rare. Growth that fast cannot last, and a company that grows too quickly often destroys itself." So the ratio is not a statement about all stocks. It is a statement about one group, and even that group has a ceiling on how fast it can grow.

Now the groups where it fails. A cyclical's P/E reads backwards: "for most stocks a low P/E is good, but cyclical stocks are not most stocks... buying a cyclical whose earnings have grown for several years at a very low P/E is an effective way to lose half your money in a short time." The same number points the other way. A turnaround is a question of survival. A slow grower is valued on its dividend. An asset play is valued on what the balance sheet does not show. Four groups, four different questions, and one ratio cannot answer all of them.

There is one more piece of support, and it comes from how Lynch himself treated the index. Running Magellan, he did not buy only growth stocks. He wrote that the fund could buy "any investment vehicle", that he did not confine himself to the style of a growth manager, and that when growth stocks in general were too expensive he moved into undervalued cyclicals. If the manager would not use a single yardstick, then testing one of his rules should start by pointing the yardstick at the thing it measures.

That is why this report deals with fast growers only, and the choice has a price, which we concede in the limitations: narrowing the scope gives up the other five groups, so nothing here can answer whether PEG works on cyclicals. What the narrowing buys is one sentence that can be said cleanly. On the kind of company Lynch himself defined as PEG's subject, does the number carry information. A narrow answer that holds beats a broad one that is vague in every clause.

What it means for you: if half of your watchlist is cyclicals and beaten-up turnarounds, "buy anything under PEG 1" does not apply to you. Not because the rule is wrong, but because it was never talking about those companies. That is the whole reason the rest of this report covers one kind of company, and the other five are outside the scope rather than left out.

Lynch is explicit in two places. Writing about Body Shop, a fast grower, his rule is: "one test of whether a stock is overpriced is that the P/E should not exceed the earnings growth rate." Writing about cyclicals, he reverses it: "For most stocks a low P/E is good, but cyclical stocks are not most stocks... buying a cyclical whose earnings have grown for several years at a very low P/E is an effective way to lose half your money in a short time." Put the two side by side and the point is that the number has a subject.

One more thing has to be said, because it decides how everything after this should be read. Lynch never screened for stocks under PEG 1 and bought them all. He visited factories, called competitors, asked suppliers whether the order book was real. PEG came last, run on a company he had already found worth checking. Treating the check as the method is treating the thermometer as the doctor.

What it means for you: if this rule is useful at all, it is useful for rejecting a name you were already considering, not for pulling twenty names out of three thousand. Those sound like neighbouring ideas. They are two entirely different methods. Which is why what follows tests the check, never the search.

What we actually tested

Testing the rule requires a decision Lynch never had to make: being explicit. He could look at a company and judge whether the story was real. A backtest cannot. It has to be told which numbers to use, how to combine them, and what counts as passing. Those decisions decide everything that follows, so they are stated here in full.

Four decisions defined the test.

Universe. Every US-listed stock, ranked by average daily dollar volume, top 3,000 kept. A five-dollar price floor removed the penny stocks. Daily prices. In one line: the most mainstream, most easily traded three thousand stocks in the US market.

Growth. This is the decision that departs furthest from Lynch, and the one that decides the sample. Earnings growth is the three-year compound rate of reported earnings at the purchase date, backed out of price and P/E. No forecasts, no analyst estimates, no research. And the rule only considers companies whose growth falls inside a fixed window, 20 to 25 percent a year. That window is not a convenience; it is the range Lynch treats as the fast grower worth holding. Everything below 20 percent is a different category with a different set of rules, and everything far above it he explicitly distrusted. So this report is not about PEG in general. It is about PEG applied to fast-growing companies, defined narrowly.

What it means for you: hold on to this one, because it is the softest spot in the report. Lynch judged what a company would do next. We measured what it had just done. Those are different variables. A company whose earnings have collapsed and is about to turn is a disaster in our data and the best idea of the year in Lynch's notebook, and that is exactly where a low PEG means something. We cannot see it.

Window. January 2012 to December 2024, thirteen years. The US market rose almost continuously through this window, with one sharp drop that recovered within months. That matters later.

Execution. One million dollars of starting capital. One dollar per order, plus 0.3% slippage on each

side, meaning the gap between the quoted price and the price actually paid. Monthly rebalancing, on the first trading day of the month, thirty minutes after the open. Equal weight per position. No position when nothing qualifies. The whole thing runs on one cloud pipeline, QuantConnect's LEAN engine with point-in-time fundamentals, so every rule sees the same data on the same schedule.

What it means for you: those costs sound trivial, a dollar and three-tenths of a percent. But rebalancing monthly means twelve round trips a year, and over thirteen years that is more than a hundred and fifty. Three-tenths of a percent compounds into a serious drag. This is why several rules below look like they should win and do not: the idea was fine, the friction ate it.

Rules. Six of them, all on the same kind of company, in three rungs. What follows from paragraph one is that PEG has a subject: the fast grower, the company whose earnings are climbing at a steady, self-funded rate. Everything else in Lynch's framework has different questions attached, and a valuation ratio divided by a growth rate cannot answer them. So the test was confined to that one type. At the bottom sits GP0, a floor far below anything Lynch would have bought: any company whose earnings grew at all over the past three years, no window, no ratio. Above it sits G, the same company with no PEG filter at all, which is the control. Above that sit the four thresholds on the ratio itself: 0.5, 1.0, 1.5, and one version that also removes the crowded growth sectors, which puts Lynch's own warning on trial.

The logic of the report sits in that paragraph, so it is worth pausing on. If PEG carries information, adding a threshold should improve results, tightening it should improve them further, and removing it should damage them. Each prediction was tested separately.

By the end the experiment had been reduced to something unusually clean. Six rules, one universe: the same 3,000 most liquid US stocks, the same monthly rebalance, the same execution and cost model. Six rules over five settings: the four thresholds on the ratio, plus one row with the ratio removed and one row with the ratio and the growth window both removed. So any difference between rows can only be assigned to those settings, with no other source.

What happened

The whole result in one sentence: six selection rules, all buying fast growers whose earnings had been climbing at 20 to 25 percent a year, produced annual returns between +3.0% and +9.3% over the thirteen years - and the one rule that dropped the growth window entirely, asking only that earnings grew at all, did better than any rule with a PEG threshold, at +9.7%. The S&P 500 returned 14.5% over the same period. The equal-weighted version of it returned 12.5%.

Put differently: the best of the six earned about five percentage points a year less than simply buying a basket of large American companies and never choosing anything. Not one of the six caught up to the S&P 500 at any point in thirteen years. The ranking, moreover, runs the wrong way: among the

rules that ask for something, the less the rule asks, the better it does - and asking for nothing at all does better than any of them.

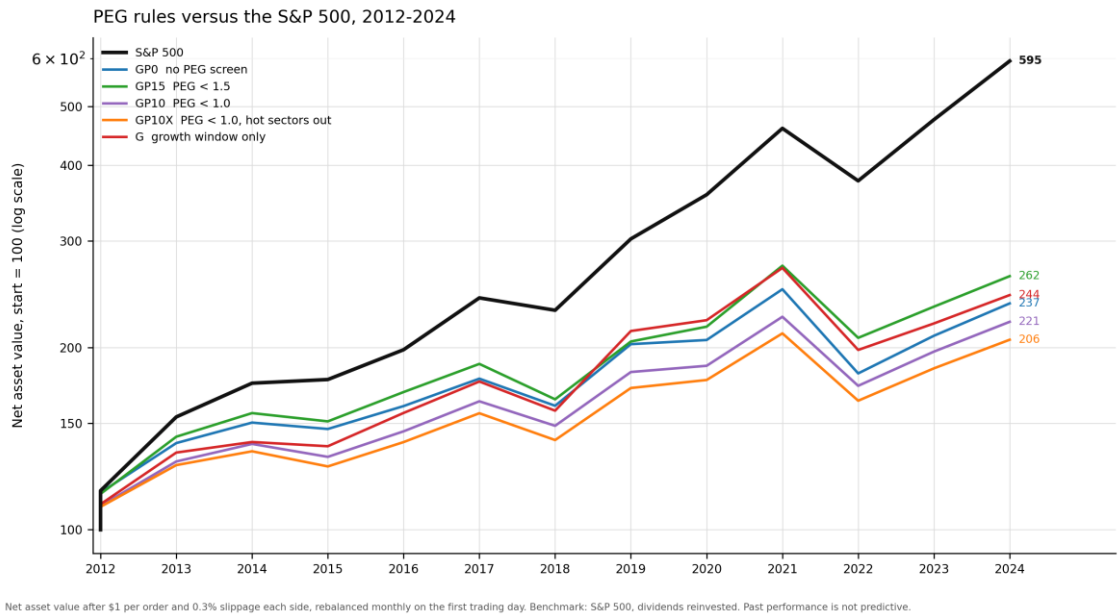
And the failure is not where a reader expects it. Tightening the gate does not steadily destroy returns; it turns a portfolio into a handful of bets. At 0.5 the rule holds six names on average, and the result is not a small loss but a wildly unstable ride: the largest annual swing in the table, 22.8%, and a correlation with the S&P 500 of 0.38. Its +3.0% over thirteen years is not distinguishable from zero. The strictest reading of Lynch's rule was the least reliable way to own the market, not the most expensive one.

Rule	What it selects	Annual	Max DD	2nd DD	Sharpe
GP0	Any growth above zero, no window	+9.70%	-39.10%	-26.90%	0.79
G	Growth window, no PEG filter	+7.10%	-39.60%	-36.70%	0.55
GP05	Growth window, PEG below 0.5	+3.00%	-68.00%	-34.40%	0.13
GP10	Growth window, PEG below 1.0	+8.80%	-49.70%	-33.30%	0.51
GP10X	PEG below 1.0, hot sectors removed	+8.50%	-55.50%	-30.20%	0.50
GP15	Growth window, PEG below 1.5	+9.30%	-47.80%	-32.50%	0.52

Three columns need translating. "Annual" is the yearly return, compounded over thirteen years. "Max DD" is how far the portfolio fell from its previous peak, on daily net asset value from 2012 to 2024. "Sharpe" is return per unit of the bumpiness you endured; 0.5 is ordinary, above 1 is good, negative means you carried risk and lost money for it.

The two drawdown columns answer different questions. The first is where each portfolio fell deepest. For five of the six that moment is the same: the February-to-March 2020 crash, troughing on 23 March. So the column records a pandemic, not a difference between rules. GP05 is the exception: it peaked on 19 February 2015, fell 68.0% into the March 2020 panic, and has not regained that peak since. A high it never gets back to is a different failure from a fall it recovers, and the annual return cannot show it.

Figure 1. Cumulative return of the six rules against the S&P 500, 2012-2024.



That is why there is a second column. It is the next-deepest fall on the same curve. None of the six had it in the pandemic, because 2020 was already taken. GP0 and G fell 26.9% and 36.7% from their late-2021 peaks into the 2022 bear market; GP10, GP10X and GP15 bottomed as late as March 2023. That column measures the rules, not the calendar.

What it means for you: there is no jackpot in this table. The best rule made 9.7% a year; the index made 14.5%. The rule closest to the index was the one that screened for almost nothing, and it still finished five points a year behind. Ten thousand dollars became about thirty-three thousand. The same money in the index became about fifty-eight thousand. More work, rebalanced every month, less money.

Read the return column again, because it is the whole experiment. Six rules, four thresholds, one floor, one control. If the ratio carried information, returns should have moved with the threshold, in one direction. They did not. Loosening from 1.0 to 1.5 gained 0.5 points; tightening from 1.0 to 0.5 lost 5.8; dropping the ratio altogether loses only 1.7 against the tightest gate. And GP0, which shares the universe, the rebalance and the costs with all of them and differs only in not screening on the ratio, beats every rule that does. Moving the threshold inside this universe produces one damaged version after another; removing it restores the best one. So the ratio is not what failed. The settings around it did: the growth window is fixed at 20 to 25 percent first, and cheapness is chosen inside that window afterwards. The first gate keeps companies that were not cheap. The second gate starts too late.

The other five did not separate either. Excluding the tightest, they finished within 2.5 percentage points of each other - about the width of one good year for a portfolio of a few dozen names. Five versions of one rule agreeing with each other means the threshold was not what decided the answer. The shared floor was.

What it means for you: in one line, PEG did not beat the market. It did not lose your money

either - it just made none that buying the index would not have made, after thirteen years of monthly rebalancing and the same swings. What it is still good for is talking you out of a name you already liked. As a way of choosing, it chose about as well as not choosing.

Two rows deserve a warning, and neither is about PEG. GP05, the tightest threshold, holds between 2.8 and 13.2 names, averaging 6.0; in 2012 to 2017 it averaged under five, and its +62.4% in 2016 came off 2.8 names. Those holdings were nearly a different set every year, which is why its second-deepest drawdown, 34.4%, came early and closed fast. That curve is a handful of positions being re-dealt, not a strategy being tested. It is a verdict on position counts that low, not on PEG.

Does PEG predict anything

Everything above compares whole strategies, and a strategy mixes several things at once: the universe, the timing, the number of holdings, the style of the year. If GP10 underperformed, PEG may have failed, or the execution may have been expensive, or the universe may have been wrong. To put the ratio itself on trial you need a cleaner question.

The cleaner question is this. On a given day, line up every eligible stock by PEG and ask whether the cheap end goes on to beat the expensive end. No portfolio, no rebalancing, no trading costs. Just: does this ordering carry information about what comes next?

We ran it twice, independently, on different windows and different code, so that the two runs could check each other.

Test	Window	Observations	IC	t value	Verdict
V14	2015-2024	21,257	-0.0086	-0.59	no predictive power
V16	2010-2024	26,485	-0.0122	-0.60	no predictive power

IC is the correlation between the ordering and the next period's return. One would mean the ordering is always right. Zero would mean it is a coin flip. Negative means it is reliably backwards. "Observations" is stocks multiplied by days, more than twenty thousand of them, so the sample is not small. The t value is the ruler for whether a result could be luck; by convention it needs to clear 2. Ours are -0.59 and -0.60, nowhere near.

In plain terms: whether a stock’s PEG is cheap or expensive tells you essentially nothing about what it does next, and whatever faint signal there is points the wrong way.

There is one more test. Split the universe into ten buckets by PEG, cheapest to most expensive, and

measure how much each bucket beat or lagged the market. A working factor should produce a smooth staircase with the cheap end on top. What we got jumps around: the cheapest decile returned 0.23% excess, the fourth 0.59%, the fifth -0.22%, the sixth 0.16%, the seventh -0.85%, the eighth 0.30%. There is no staircase. If you had to guess from that pattern which bucket did best, you could not.

What it means for you: this is the hardest evidence in the report, because no excuse about fees, universe or rebalancing touches it. No trading is involved. It says only that in this market, over this period, the ratio has no measurable ordering power. Sort by it and the order you get is not meaningfully different from random.

Can nobody else reproduce it either

The first suspicion a negative result should invite is that the test was built wrong. So we went looking for the opposite result. Not a description of Lynch's method, but someone who ran a mechanical version and reported what it earned.

We found four independent attempts spanning twenty-five years, plus this project. The figures are taken from each source's own public disclosures, and each is compared with its own stated benchmark.

Source	What it is	Period	Annual	vs benchmark
SPGP	Real fund, $2.1bn, S&P 500 GARP	2011-2026	14.15%	-0.30pp
tradingstudio	Public backtest, US market	2000-2025	7.10%	-0.91pp
finlab	Public backtest, S&P 500	2018-2026	11.60%	-3.50pp
This project	Backtest, GP10	2012-2024	8.80%	-5.70pp

The last column is the only one that matters here: return relative to the matching market index over the same period. Every one is negative. The best is a dead heat.

One clarification, so the table is not read as a scorecard on other people. The point is not whether these funds or authors did their work well. They appear for one purpose: to establish that our shortfall is not bad coding, because the same approach in other hands, with real money and different methods, points the same way. When five independent attempts converge, the likelier explanation is the rule.

One more thing that is easy to misread. The comparison is on the "vs benchmark" column alone,

meaning whether the rule beat its market, not on who earned more in absolute terms. Different windows and different market conditions make absolute returns incomparable.

What it means for you: if you see someone posting a spectacular PEG strategy, look for the start date and the cost assumption before the return. Those two decide most of the answer, and they are usually the hardest to find in the post.

A negative result is easiest to over-read when it is your own. This section is short on purpose. The results do not overturn Lynch, and they are not evidence that the test was built badly. They overturn something smaller.

One: the growth input is not Lynch's input. We measure the last three years of reported earnings; Lynch estimated the next three by talking to the company, its competitors and its customers. Those are not the same variable with noise on it. So the number never tested here is the one Lynch used - only a stand-in that history can compute.

Two: the window covers one regime, and the start date was not a free choice. Point-in-time fundamentals are the binding constraint: a company's reported earnings and the share count that divides them, as they were filed, before any restatement. That coverage on a 3,000-name liquid universe begins in 2012. Earlier screens draw on restated figures and on companies still listed today, which is exactly the bias the whole design was built to avoid. So the window is a compromise, not a selection. It matters in both directions. From 2012 to 2024 the US market had no real bear market; what the rule would do in 2000, in 2008, or across a long flat stretch, this report cannot say. But the window was also favourable to the rule in the way that most flatters a stock picker: a rising market, cheap money, growth stories bid up - and the rule still lost to the index. That is harder to wave aside than a loss in a bad decade would have been.

Three: turnarounds are barely tested here, and that is a consequence of the scope, not an oversight. The report covers one kind of company, the fast grower, because that is where PEG has a meaning. Nothing here says PEG fails on a company whose earnings just collapsed. It says PEG was not tested there.

What the number is good for

The useful output of this project is not that PEG is worthless. It is a clearer statement of what the number can and cannot do, measured on the one class of company where Lynch actually used it.

It cannot pick stocks. Ranked against the market on its own, it carries no measurable information about what happens next, and every portfolio that used it as its main filter underperformed. The IC tests say this, and so do the five rules; the two methods fail for different reasons.

It can still keep you honest. A PEG near 1 is a reasonable way to check whether the price you are paying for a growth story is sane. Lynch used it that way, at the end of his process rather than the

start.

What our results rule out is the use of it as a screen: the belief that sorting fast-growing companies by this ratio and buying the cheapest end is a strategy that can be automated.

The difference matters because the sort is easy to automate, and the thing that made Lynch's method work is not. His advantage was knowing which company he was looking at and why it deserved a look, which is exactly what a screen cannot tell you. What we have shown is that the automatable part of the method is not where the value was.

What it means for you: next time a story catches your eye, ask how much optimism is already inside the price rather than whether the PEG is under 1. The first question is what Lynch was actually doing. The second is a checkbox.

One thing would move this forward more than anything else: replacing trailing growth with an expectation of future growth, the input Lynch actually used. That would test the version of the rule he would recognise. It is not a small job - it is the difference between having tested a proxy and having tested the rule.

Until then, the honest summary is narrow. A rule that a large number of investors believe in, applied to the one category of company it was made for, performed no better than owning the market, and no better when the thresholds were moved in either direction.