Growth expectations in the AI era

Levels, uncertainty and upper tails in professional GDP forecasts

Max Ghenis

2026-09-19

Abstract

A stated growth distribution can move toward faster growth in three places: a higher central forecast, a wider spread, or more probability on rapid growth. We measure all three, in the US and ECB Surveys of Professional Forecasters through 2026Q3, holding the survey quarter and calendar horizon fixed. The main comparison sets 5 Q1 rounds (2015–19) against 2 (2025–26), so it describes the recent rounds and cannot establish a trend. Mean next-year US growth forecasts fell from 2.19% to 1.89%, the reconstructed pooled standard deviation rose from 1.28 to 1.44 percentage points, and the probability of growth above 4% fell from 5.62% to 4.22%. Euro-area forecasts four years ahead fell by about as much from a lower base, from 1.64% to 1.35%, while their probability of growth above 4% rose from 0.43% to 1.06% — small in absolute terms, and 2.5 times its starting value. Across levels, spreads and tails, the surveys show limited movement toward rapid growth. These facts describe the overall growth outlook inside the surveyed horizons. They do not estimate AI's effect, and they cannot show that forecasters ignore AI. We bound high-growth probabilities directly from the reported bins, keep the COVID episode in view, and separate this evidence from longer-horizon and conditional AI scenarios.

1 Where a growth transformation would show up

A median forecast near 2% does not mean a forecaster is confident growth will stay near 2%. Someone who puts a small but consequential probability on a large acceleration can report an ordinary median alongside an unusual upper tail. Comparing central estimates therefore cannot show how much probability forecasters put on rapid growth. The distributions can. They describe aggregate growth, so they cannot isolate what forecasters believe about AI: a larger expected AI contribution can sit inside a lower total.

The Surveys of Professional Forecasters supply them. Respondents attach probabilities to ranges of GDP growth, which lets us ask three questions. Have central forecasts moved toward faster growth? Have the distributions widened? Has more probability moved into high-growth outcomes? We answer all three reproducibly, across US and euro-area horizons, with twenty years of history and an explicit account of what the bins can and cannot identify.

The record shows no large upward shift. Central forecasts sit below their pre-COVID levels. Uncertainty rose around COVID and then declined, with some measures still above where they started. The high-growth tail is small in aggregate, and it has grown in the euro area. Together these say more about the expected growth outlook than the narrow dispersion of long-run point forecasts can.

What the record contrasts with depends on the claim. A forecast that puts substantial unconditional probability on rapid US or euro-area growth by the late 2020s or early 2030s meets a measurable contrast in these distributions. A scenario in which advanced AI eventually produces explosive global growth asks a different question, and we keep the two apart instead of placing unlike projections on one numerical scale.

2 Data and comparison design

The pipeline behind this paper produces 3,695 density summaries across every surveyed variable and 3,070 forecast–outcome pairs. Here we use real GDP growth. US real GDP density forecasts begin in 1992; the earlier nominal and real GNP densities stay in the tracker and out of these comparisons. ECB GDP densities begin in 1999 (Federal Reserve Bank of Philadelphia 2026; Allayioti et al. 2024).

We follow three series: the US distributions for the next calendar year and for three calendar years ahead, and the ECB longer-term distribution. We use Q1 rounds throughout. That fixes the calendar horizon and makes the ECB longer-term series a forecast four years ahead; in Q3 and Q4 the same question refers to a year five years ahead. Each series forecasts annual growth in its target year, and none averages growth over the intervening years.

The main comparison sets 2015–19 against 2025–26: 5 pre-COVID Q1 rounds against 2 recent ones. The figures show the full path from 2006 where the series exist (US year +3 starts later) and keep 2020–24, a window that holds both the pandemic and the US wide-bin regime of 2020Q2–2024Q1. The US windows also sit in different bin schemes: one-point bins through 2020Q1, and a coarser grid since 2024Q2. Coarser bins raise a reconstructed standard deviation mechanically, so part of the US increase reflects the questionnaire. The euro-area bins are identical in both windows.

The companion reports 2010–19, 2010–14, 2020–24 and 2023–26 windows and an all-round sensitivity that averages within each survey year before averaging across years, so the three rounds available in 2026 weigh as much as a full year. All-round comparisons still mix horizon lengths within the year. The panel changes between rounds, so nothing here identifies within-person updating or a causal effect of generative AI. We run no equivalence test, and two recent rounds cannot prove that the process is unchanged.

3 Central forecasts and stated uncertainty

Table 1 compares means and spreads at fixed horizons. The standard deviation and the interquartile range both describe the pooled predictive distribution. Neither measures how far forecasters' point forecasts sit from each other.

Table 1: Q1 forecast levels and spreads, 2015–19 (five rounds) against 2025–26 (two rounds). Each cell shows earlier ⟶ recent; means are percentages and spreads are percentage points.

Forecast	Mean	Pooled SD	Pooled IQR
US, next year	2.19 ⟶ 1.89	1.28 ⟶ 1.44	1.45 ⟶ 1.54
US, year +3	2.02 ⟶ 1.91	1.42 ⟶ 1.47	1.67 ⟶ 1.46
Euro area, year +4	1.64 ⟶ 1.35	0.83 ⟶ 0.90	0.97 ⟶ 0.95

Mean growth forecasts fall in all three comparisons. The US next-year and euro-area falls are close in percentage points, and the euro-area fall is the largest in proportion to its starting level. The reconstructed widths do not show a radically broader range of outcomes. The US next-year standard deviation rises moderately, partly through the bin change; the three-year-ahead and ECB four-year-ahead standard deviations move less, and the interquartile ranges at those two horizons narrow slightly. Limited movement toward faster and more uncertain growth describes these rounds. "Unchanged" does not.

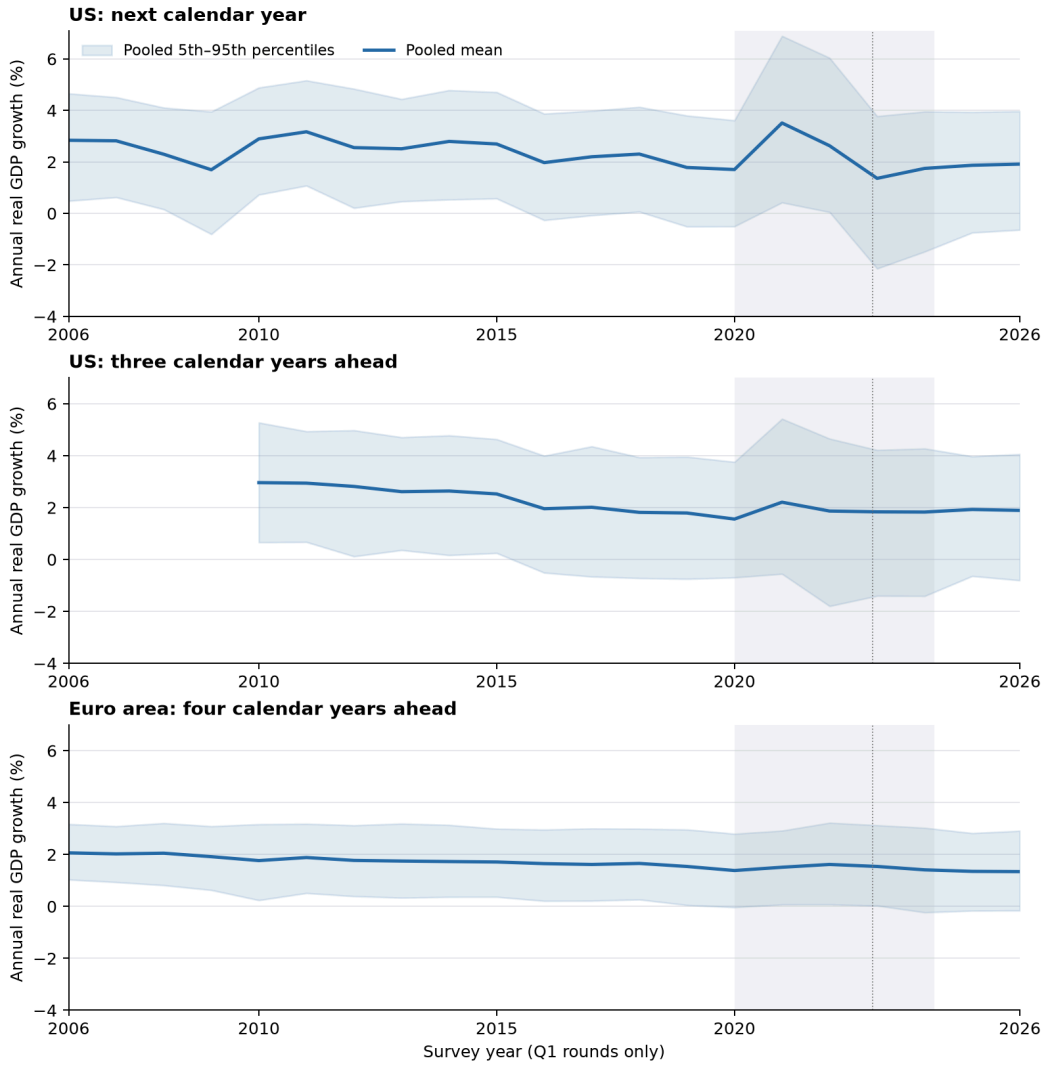

Figure 1: GDP forecast means and reconstructed pooled 90% intervals, Q1 rounds. Gray shading marks 2020–24; the dotted line marks late 2022 as a visual reference and carries no treatment interpretation. The US year +3 question arrived in 2009Q2, so its Q1 series starts in 2010.

The longer record also rules out calling central forecasts constant. The US ten-year median point forecast fell from 3.2% in 2006 to 2.1% in 2026. That question asks for average annual growth over a decade and collects a point forecast. The spread of those answers measures disagreement about a point estimate, and no predictive interval stands behind it. We report its level separately and draw no term structure of uncertainty from it.

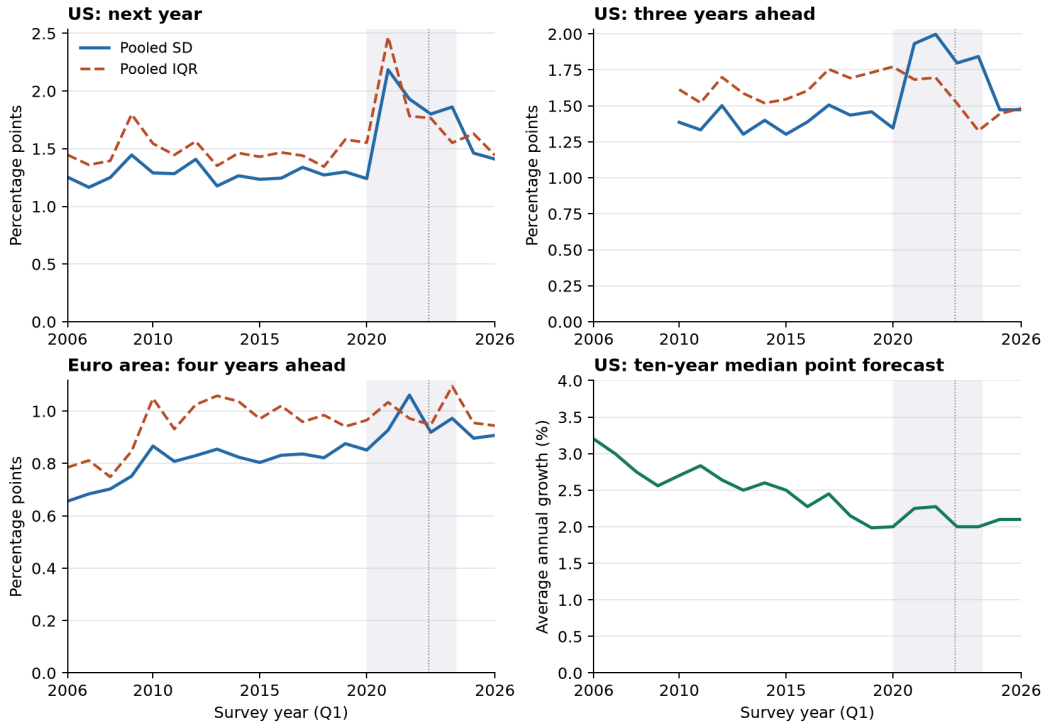

Figure 2: Predictive SD and IQR at three fixed horizons, alongside the US ten-year median point forecast. The final panel shows a different outcome and contains no elicited predictive interval.

4 The upper tail

We compute probabilities of growth above 3%, 4%, 5% and 10%. The thresholds describe outcomes and define nothing about transformative AI. For historical comparison 4% does the most work: it is a printed lower bin label in every scheme behind the displayed series, so its exceedance probability needs no assumption about how mass spreads inside a bin. It does take the printed label as the boundary; the method section reports what happens if the boundary sits at 3.95 instead.

Table 2: High-growth probabilities from the bin edges, Q1 rounds, 2015–19 (five rounds) against 2025–26 (two rounds). Ranges are identification bounds set by bin resolution; they are not sampling confidence intervals.

Forecast	Period	Above 4%	Above 5%	Above 10%
US, next year	2015–19	5.62%	1.19%	0.00–0.32%
US, next year	2025–26	4.22%	0.52–4.22%	0.00–0.05%
US, year +3	2015–19	5.97%	1.69%	0.00–0.45%
US, year +3	2025–26	4.85%	0.85–4.85%	0.00–0.07%
Euro area, year +4	2015–19	0.43%	0.00–0.43%	0.00–0.43%
Euro area, year +4	2025–26	1.06%	0.00–1.06%	0.00–1.06%

The US probability above 4% falls at both horizons between the five earlier rounds and the two recent ones. The ECB's four-year-ahead probability rises from 0.43% to 1.06%: 2.5 times its starting value, and still about one percent. A summary that says the upper tail moved nowhere would miss it.

The higher thresholds show what the bins cannot say. In the ECB scheme, all growth at or above 4% shares one open bin. Its entire probability could lie below 5%, above 10%, or anywhere between, so we

can bound those probabilities and cannot separate them. The US scheme in force since 2024Q2 places 5% inside a finite bin, which widens the identified range relative to the earlier scheme. Fitting a distribution inside those bins would add precision from an assumption, with no new survey response behind it.

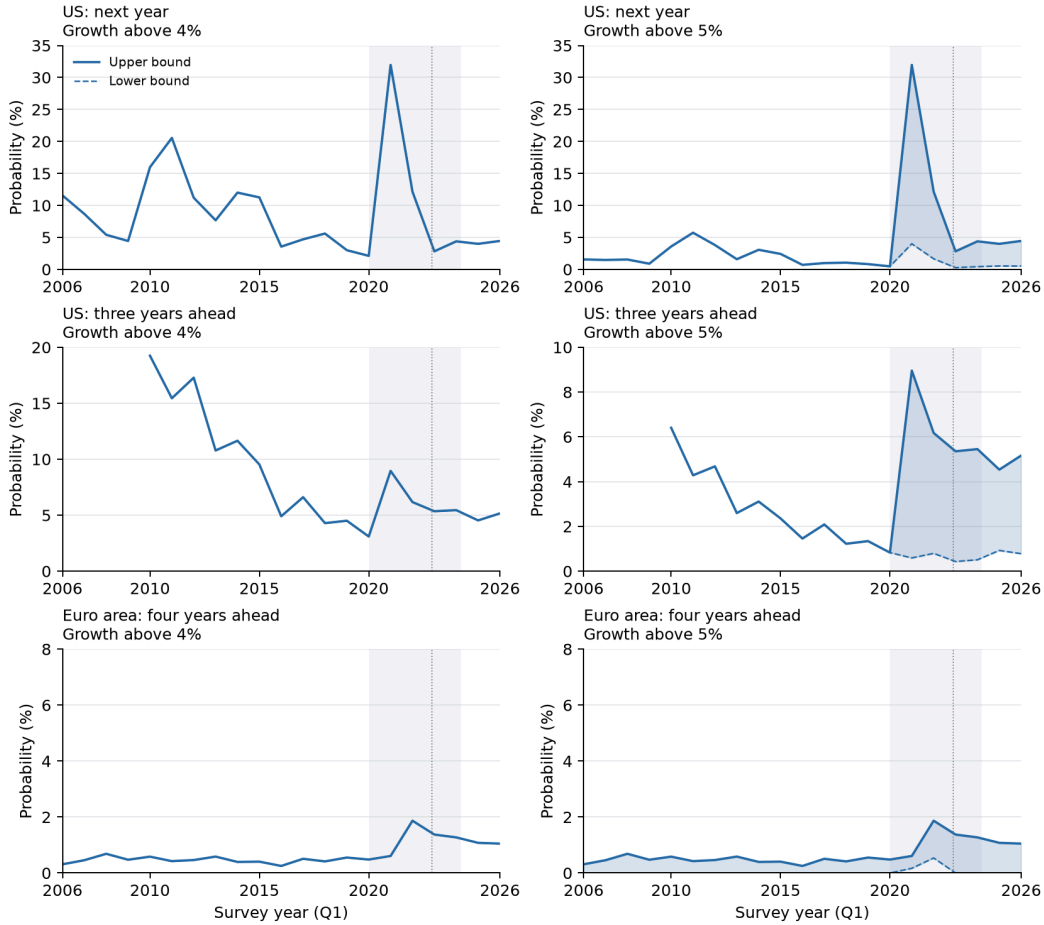

Figure 3: Upper and lower bounds on growth probabilities, Q1 rounds. Where the bounds coincide, a bin edge identifies the probability. The shaded range reflects bin resolution and says nothing about sampling error.

The latest round tells the same story. In 2026Q3 the pooled probability above 4% is 3.51% for next-year US growth and 1.45% for ECB longer-term growth, which in that round targets 2031. In the ECB round, 10 of 27 retained respondents put positive probability in that bin. A small pooled tail leaves room for many respondents who do not rule out high growth.

5 What AI growth forecasts say, and what they can be compared with

The closest comparison uses elicited distributions, because similar medians can hide different tails. Karger et al. (2026) elicit beliefs about US GDP growth as five-year annual averages. Their Figure 5 plots pooled distributions for 2030 and 2050 and labels the share of each distribution above 10%. For AI experts the 2030 panel shows 0.0% unconditionally and 3.5% conditional on rapid AI progress; the 2050 panel shows 1.4% and 10.4%. The plotted values are rounded. This source does not establish a large unconditional near-term disagreement with the SPF in the extreme tail. The averaging interval, the elicitation and the population all differ, and a conditional distribution needs scenario probabilities before it can sit beside an unconditional one.

Other prominent estimates differ more basically, in outcome or in horizon:

Source	Outcome and interpretation	Relation to this paper
Acemoglu (2024)	Task-based estimate of AI's contribution to productivity over ten years	An incremental productivity effect, with no probability distribution for annual GDP growth
Davidson (2021)	Possibility of global growth above 30% per year before 2100	Shows why extreme outcomes matter; lies beyond the SPF's geography and horizon
Erdil and Besiroglu (2023)	Arguments for explosive growth from broad AI automation	Ties automation assumptions to growth regimes; supplies no matched annual US probability

The SPF evidence contributes something specific. The professional outlook has not shifted toward substantially faster growth in the center, and its reported upper-tail mass puts numbers on high-growth outcomes inside its horizons. Beliefs that assign appreciable unconditional probability to those outcomes over the same horizons would contrast with it. The sources above do not yield one commensurable probability difference between two populations, and we do not build one by treating global, conditional or century-scale scenarios as US annual forecasts.

The outlook could stay broadly similar while respondents revise their views about AI. Positive AI effects could offset demographic or other headwinds. Respondents could expect deployment after the forecast horizon. Some could revise up while others revise down. Bins and rounding could hide changes in very small probabilities. The data establish none of these; each needs matched questions and respondent-level evidence on AI assumptions.

6 How much weight the stated probabilities deserve

The historical record helps decide how much weight the stated probabilities deserve. It cannot say which view of AI is right. For Q1 next-year US GDP, a band of one pooled standard deviation around the mean covers 22 of 33 outcomes. The pooled 90% interval covers 26 of 33 (78.8%), below its nominal coverage. One-standard-deviation coverage does not test calibration for a non-Gaussian distribution. The misses include the late-1990s growth acceleration, the 2008–09 recession and the pandemic.

We also score whole distributions with the continuous ranked probability score (CRPS) (Gneiting and Raftery 2007). For each benchmark, skill equals one minus mean forecast CRPS divided by mean benchmark CRPS on the same eligible observations, which compares average loss directly. Averaging per-observation percentage skill weights outcomes differently, so we keep it as a sensitivity and draw no conclusion from its sign.

Table 3: GDP CRPS skill against expanding benchmarks, all scored horizons. Positive values favor the survey. Each benchmark uses its own matched eligible sample.

Series	Benchmark	Rows	Ratio of mean CRPS	Mean row skill
US GDP	climatology	382	38.5%	41.2%
US GDP	gaussian	312	6.3%	5.5%
Euro-area GDP	climatology	489	19.9%	16.6%
Euro-area GDP	gaussian	480	1.2%	17.2%

Forecast–outcome pairs can share a realized target, and rolling outcomes overlap, so these summaries describe the record and establish no statistical significance. The Gaussian benchmark learns the RMSE of prior forecast errors; climatology uses prior outcomes. Both require ten historical observations, and a target period must end before the forecast quarter to count. We do not reconstruct historical publication dates or data vintages. The evaluation uses an expanding window with revised outcomes and makes no claim to be a real-time backtest.

7 Measurement and reproducibility

Within a round we keep respondent histograms whose probabilities lie between 0 and 100 and sum to within two percentage points of 100. Partial missing entries count as zero, and entirely missing histograms drop out. We normalize retained probabilities and weight respondents equally. The filter rejects two negative ECB entries in 2023Q1 along with responses whose totals are invalid.

The questionnaires print closed one-decimal ranges ("2.0 to 2.9", then "3.0 to 3.9"). We read each range as running to the next range's lower label, so "2.0 to 2.9" covers [2, 3) and lower labels such as 4.0 stay bin edges. For moments, quantiles and scores, probability spreads uniformly inside each bin, and an open tail receives the width of its nearest finite bin. Reading the printed labels literally instead lowers each mean reported here by 0.05 points and leaves each comparison between periods unchanged (`outputs/reconstruction_sensitivity.csv`). If bin b has midpoint x_b and width w_b, respondent i has

$$\mu_i = \sum_b p_{ib} x_b, \qquad v_i = \sum_b p_{ib} \left((x_b - \mu_i)^2 + \frac{w_b^2}{12} \right).$$

The equal-weight empirical mixture has variance

$$V = \frac{1}{n} \sum_i v_i + \frac{1}{n} \sum_i (\mu_i - \bar{\mu})^2.$$

The first component is average individual uncertainty and the second is disagreement (Wallis 2005; Zarnowitz and Lambros 1987). The denominator is n because the object is the finite empirical pool; a sample-variance correction would estimate a different object. Numerical-integration tests check the identity against the pooled distribution. The reported within SD is the square root of average individual variance, which differs from the average individual SD.

For high-growth probabilities we drop the finite-tail closure. When a threshold falls inside a bin, the bin's entire mass counts toward the upper bound and none toward the lower bound. Bins entirely above the threshold count toward both. At an exact edge we assume continuous outcomes, so the boundary itself carries no mass. Averaging respondent bounds gives pooled bounds. The pipeline also emits a uniform interpolation unless the threshold lies strictly inside an open tail; the headline bounds never use it. Small mass in an open tail does not bound how fast growth is, conditional on reaching that tail. The bins alone therefore identify no unrestricted mean or variance: the level and width results depend on the reconstruction. The exceedance bounds use neither the tail closure nor any within-bin shape, but they do depend on where a bin's boundary sits. Under our reading, and with the printed labels taken literally, 4.0 is a boundary and the probability above it is a single number. If outcomes instead round to one decimal, the boundary sits at 3.95 and 4% falls inside a bin. The US next-year probability is then bounded at 1.19–5.62% for 2015–19 and 0.52–4.22% for 2025–26, and the euro-area four-year probability at 0.00–0.43% and 0.00–1.06%. Those ranges overlap, so under that reading the surveys establish neither the US fall nor the euro-area rise in the probability of growth above 4%.

These distinctions matter across survey redesigns. Neither the SD nor the IQR is invariant to changing bins, and probabilities at a common bin edge give the cleanest historical tail comparison. Respondent rounding also limits very small probabilities: zero reported mass records the elicited answer and cannot prove that a respondent's latent probability equals zero.

Annual US outcomes use the archived BEA series. Annual ECB GDP outcomes use the ECB's published annual real GDP growth rate, retrieved on 17 September 2026. Calendar inflation uses the official annual-average index-growth series, published to 0.1 percentage point and retrieved on 5 September 2026. Rolling observations keep the archived files, whose retrieval dates were not recorded. Each annual outcome carries the union of the official annual status flag and the archived subannual flags for that year. `data/raw/ecb_annual_realizations_sources.json` records the URLs, hashes, precision and verification checks.

The source repository holds the raw snapshots, parsers, numerical tests, `growth_tails.csv`, the period comparisons, score summaries, figure code and this manuscript. Rebuilding runs the pipeline build, the research figure build, the tracker data build and the manuscript render. The growth companion offers baseline, round and threshold controls over the same series.

8 Conclusion

Forecast levels, widths and tails tell one story together. In the two recent Q1 rounds, as in the five before COVID, the professional outlook in these surveys centers on modest growth, and little probability has moved toward much faster outcomes at the available horizons. Looking past point estimates does not change that. One thing has moved: the ECB's small longer-term high-growth tail has grown.

The record gives near-term growth beliefs a concrete professional baseline. It cannot show that respondents have neglected AI, and it cannot resolve projections beyond the surveyed horizons. A direct comparison would ask professional macroeconomic forecasters and AI specialists for distributions over the same US annual and multi-year growth outcomes, on the same dates, both unconditionally and conditional on stated AI capabilities. The existing surveys supply the baseline for that comparison.

Bibliography

Acemoglu, Daron. 2024. *The Simple Macroeconomics of AI.* Working Paper No. 32487.

Allayioti, Anastasia, Rodolfo Arioli, Colm Bates, et al. 2024. *A Look Back at 25 Years of the ECB SPF.* Occasional Paper Series No. 364.

Davidson, Tom. 2021. "Could Advanced AI Drive Explosive Economic Growth?.". https://coefficientgiving.org/research/could-advanced-ai-drive-explosive-economic-growth/.

Erdil, Ege, and Tamay Besiroglu. 2023. "Explosive Growth from AI Automation: A Review of the Arguments.".

Federal Reserve Bank of Philadelphia. 2026. "Survey of Professional Forecasters: Documentation.".

Gneiting, Tilmann, and Adrian E. Raftery. 2007. "Strictly Proper Scoring Rules, Prediction, And Estimation." *Journal of the American Statistical Association* 102 (477): 359–78. https://doi.org/10.1198/016214506000001437.

Karger, Ezra, Otto Kuusela, Jason Abaluck, et al. 2026. *Forecasting the Economic Effects of AI.* Working Paper No. 35046. https://forecastingresearch.org/research/economic-effects-of-ai.

Wallis, Kenneth F. 2005. "Combining Density and Interval Forecasts: A Modest Proposal." *Oxford Bulletin of Economics and Statistics* 67 (s1): 983–94.

Zarnowitz, Victor, and Louis A. Lambros. 1987. "Consensus and Uncertainty in Economic Prediction." *Journal of Political Economy* 95 (3): 591–621.