VULTAX RESEARCH NOTES 2026-035

What the Market Expected: 153 Historical Economic-Event Auctions

Christopher Geneth Kurg, Vultax Research, Tallinn

ORCID https://orcid.org/0009-0007-4902-8847

PUBLISHED 25 October 2026

PAGE https://vultax.com/research/economic-derivatives-auction-archive

DOI No DOI has been assigned.

ABSTRACT

Explore 153 historical economic-release auctions: probability ranges, announced outcomes, source lineage and credits to Gürkaynak and Wolfers.

CITE AS Kurg, Christopher Geneth (2026). What the Market Expected: 153 Historical Economic-Event Auctions. Vultax Research notes 2026-035, Vultax Research, Tallinn, 25 October 2026. https://vultax.com/research/economic-derivatives-auction-archive

AI-assisted drafting, research and editing. Published by OmniOS OÜ.

VULTAX RESEARCH · ECONOMIC DERIVATIVES

What the market expected

Before a US economic release, an auction priced every range of possible answers. Refet Gürkaynak and Justin Wolfers collected 153 of those distributions. Explore them, and see what a price can and cannot tell you.

Published by Christopher Geneth Kurg · Vultax Research, Tallinn · 25 October 2026

> **Original research and auction data: Refet S. Gürkaynak and Justin Wolfers.** This reading edition and explorer were prepared by Vultax Research with Justin Wolfers's permission to republish the auction data. The authors have not reviewed it, and it implies no endorsement.

On 9 January 2004, before the US government reported December's change in nonfarm payroll employment, traders in a short auction put a price on every plausible answer.

Almost half of their probability sat between 100,000 and 200,000 new jobs. Before you reveal what was announced, look at how much room they left for a surprise.

A price on every outcome

In October 2002, Goldman Sachs and Deutsche Bank began running auctions tied directly to upcoming US data releases. The simplest contract to think about is a digital range: it pays $1 if the announced number lands between two adjacent strike prices, say payrolls between 100,000 and 125,000, and nothing otherwise. If traders were indifferent to risk, the price of each range would be the market's probability that the number lands there. (Gürkaynak and Wolfers, 2005, pp. 1–3)

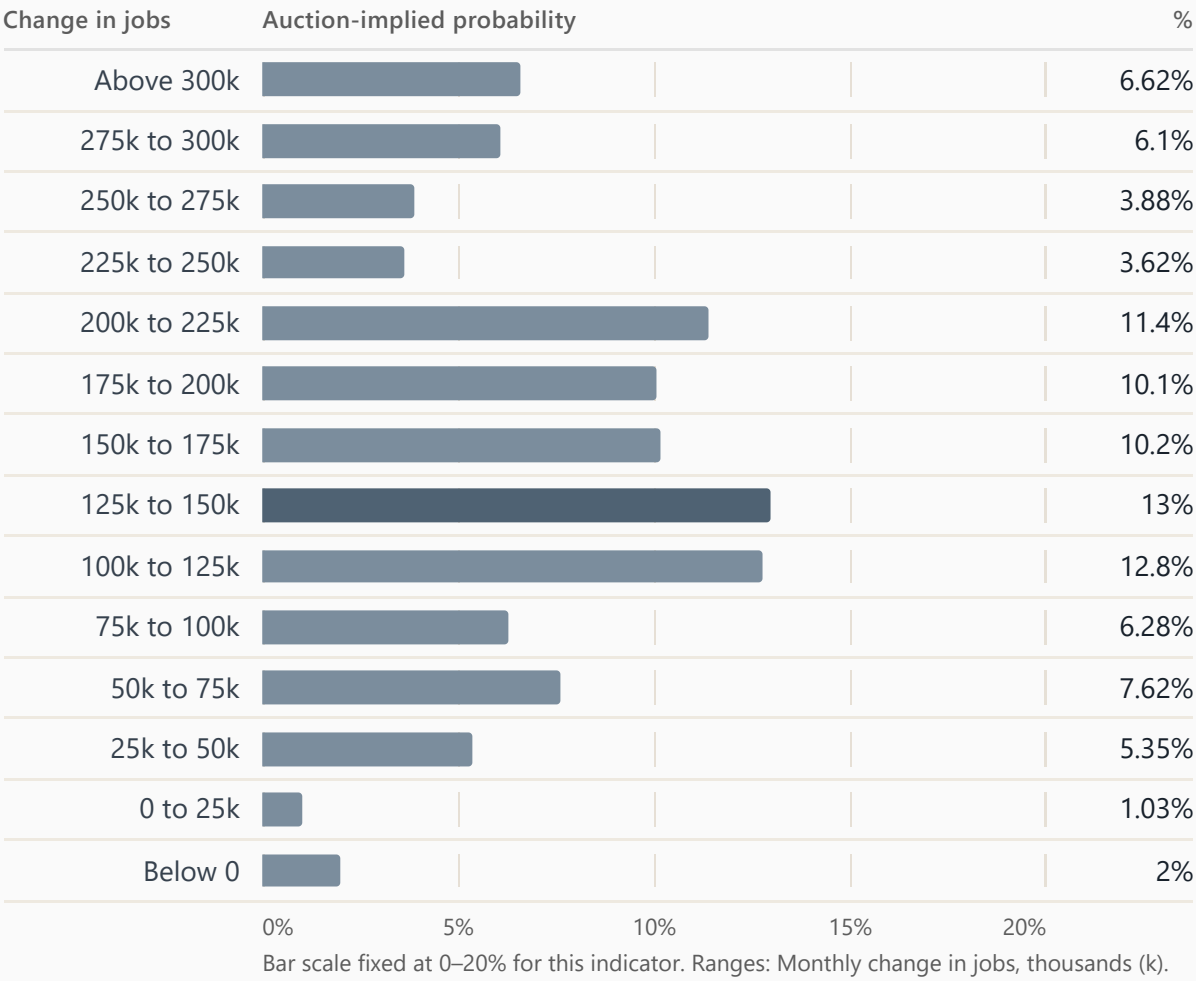
The auctions worked like a racetrack pool: money on every range was pooled and paid to the range that happened, and final prices were known only when an auction closed, usually after about an hour on the morning of the release. Payrolls traded most actively, jobless claims least. (pp. 3–5)

Gürkaynak and Wolfers assembled the first two and a half years of results: 153 releases between October 2002 and June 2005, across four indicators. There are 33 monthly payroll reports, 30 readings of the ISM manufacturing index, 26 releases of retail sales excluding autos and 64 weekly counts of initial jobless claims. Each release has between 10 and 18 priced ranges, 2,235 in all. Figure 1 starts with the December 2003 jobs report.

FIGURE 1 · A MARKET'S ODDS ON ONE NUMBER

Nonfarm payrolls, released 9 Jan 2004

Auction 9 Jan 2004 · 14 ranges · published probabilities total 100.00%



Bar scale fixed at 0–20% for this indicator. Ranges: Monthly change in jobs, thousands (k).

Most heavily priced range: 125k to 150k, 13%.

Announced figure hidden. Look at where the auction put its weight, then reveal it.

▶ **Source, units and citation for this release**

Figure 1. Auction-implied probability for each range of the selected indicator, as published in the Gürkaynak–Wolfers data. Probabilities are not rescaled; the lowest and highest ranges are open-ended. Reveal the announced figure, switch between published probability and probability per unit of width, then browse all 153 releases.

The announced change in nonfarm payrolls was 1,000 jobs: one thousand, across all of US nonfarm payroll employment. The two lowest ranges, below zero and zero to 25,000, together carried 3.03 percent. That is a price on any result below 25,000, not on exactly 1,000.

By the market's own pricing, a result that low was roughly a one-in-thirty event. That does not make the market wrong. If forecasts like this are calibrated, about three in every hundred three-percent events should happen, and a single case cannot tell you whether this was one of them. It is also why we led with it: an example chosen because it is surprising says nothing about how often surprises occurred.

An ordinary month, and a tie

To show what the archive looks like when nothing dramatic happens, we fixed a rule before looking further: take the first payroll release after January 2004 whose announced figure fell inside the range holding the market's midpoint, the point with half the published probability below it. The rule lands on the December 2004 report, released on 7 January 2005. The announcement was 157,000, inside the 150,000–175,000 range. Ranges below it held 49.0 percent.

Notice what "ordinary" means here. Even the range containing the market's midpoint carried only 7.5 percent. A wide distribution spreads its probability across many ranges, so the price of the range that happened is not, on its own, a measure of surprise. Where the outcome sits in the cumulative distribution is more informative: in this case, between 49.0 and 56.5 percent of the probability lay below it.

A second rule picks a case this edition leaves unresolved. Twenty-three announcements fall exactly on a range boundary, 14 of them jobless-claims releases with ranges 5,000 claims wide. The earliest of the 14 claims cases is 26 February 2004: 350,000 claims, precisely between 345,000–350,000 (14.77 percent) and 350,000–355,000 (14.17 percent). The ranges below that line carried 63.28 percent of the published probability. Whether that equals the chance of a lower figure, and which range paid out, depended on how the contracts settled a tie, which the package does not record, so the edition assigns the result to neither side.

Surprise December 2003 payrolls, released 9 Jan 2004. Chosen for its surprise.

Ordinary December 2004 payrolls, released 7 Jan 2005. Chosen by the midpoint rule.

Tie Initial claims, 26 Feb 2004. Earliest claims result on a boundary.

Across all 153 releases, 116 announcements landed inside a finite range, 14 in one of the open-ended ranges at either end and 23 on a boundary. Those counts describe the archive. They are not a score, and three hand-picked examples cannot stand in for one.

Reading a bar: mass, not density

Each bar shows probability mass, the published price of a whole range. When ranges have different widths, the tallest bar is not necessarily where outcomes were thought most likely to cluster. The first payroll auction in the archive, for the September 2002 report, mixes ranges 10,000, 20,000, 30,000 and 50,000 jobs wide. Its tallest bar, 11.56 percent on 50,000–70,000, is not its most concentrated: 8.67 percent sits on 0–10,000, a range half as wide. Seventeen auctions, 12 ISM and five payroll, mix widths like this.

Switch Figure 1 to “per unit of width” to see the difference. That view divides each finite range’s published probability by its width. The open-ended ranges have no width, so they get no bar. It is a derived view for reading shape, not a source value, and it assumes nothing about how probability is spread inside a range. Open the September 2002 report in that view.

Widths also change between indicators and over time, so the tallest bars of two auctions do not compare their confidence. Totals need care too: 32 releases miss 100 percent by more than 0.05 points, by at most 0.6. They are shown as published.

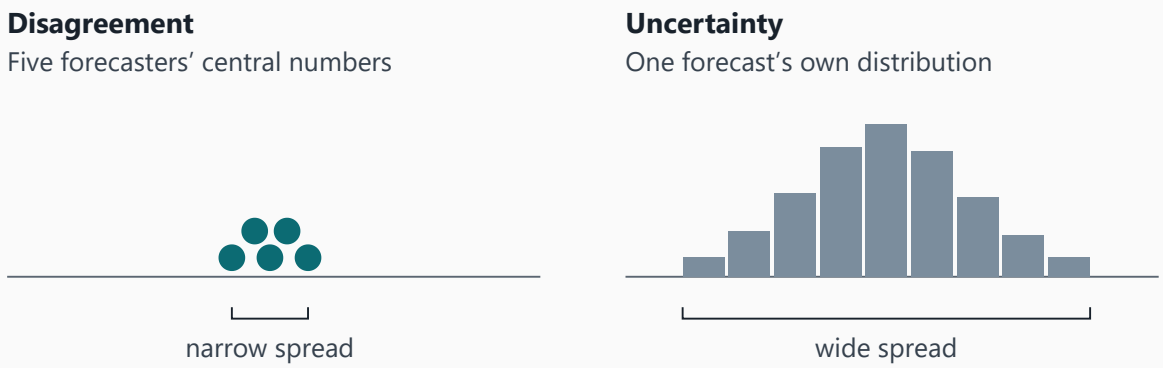
Uncertainty is not disagreement

A full distribution measures something economists rarely observe: how uncertain a forecast was. Gürkaynak and Wolfers use the standard deviation of each auction’s distribution as their measure of uncertainty and set it against disagreement, the spread of point forecasts among the professional economists in a survey. The two are different. Forecasters can agree closely on a central number while each is quite unsure of it, and the authors find exactly that pattern: release by release, disagreement is a poor stand-in for uncertainty. (pp. 2, 17–18)

FIGURE 2 · TWO DIFFERENT SPREADS

Disagreement between forecasters, uncertainty within a forecast

ILLUSTRATION, NOT SOURCE DATA



Both panels share one horizontal axis: possible values of the release.

Figure 2. Forecasters can agree closely on a central number (first panel) while any one forecast still spreads its probability widely (second panel). The survey data behind the authors' disagreement measure come from other providers and are not part of this edition, so this figure is a schematic, not a plot of them.

The authors also set the market's sense of its own uncertainty against the size of its errors. For payrolls, the average market-implied standard deviation was 96,100 jobs; the standard deviation of its forecast errors was 100,700 (Table 4, p. 17). Any standard deviation from ranges needs assumptions about where probability sits inside each range and how far the open tails reach; the paper uses range midpoints and states its tail treatment (p. 6). This edition publishes the ranges and leaves those choices to the analyst.

Prices are not quite beliefs

Reading a price as a probability assumes traders were indifferent to risk. A contract that pays when the economy disappoints is a kind of insurance, and insurance can carry a premium that lifts its price above the true probability. Gürkaynak and Wolfers conclude that for these release-day contracts the premium is small enough to ignore for many purposes (section 6, from p. 30). That is their finding under their assumptions; here, probability means auction-implied probability, and we have not tested the question independently.

The number on the day

The contracts paid on the first official estimate, not the figure statisticians settle on years later (p. 5), and this edition keeps that announced value. The difference matters. Today's Bureau of Labor Statistics series puts December 2003 payroll growth at about 109,000 jobs, inside one of the two ranges the auction priced most heavily. The +1,000 was the number on the day, and the number the auction was about. The announced 157,000 for December 2004 is now about 127,000. Check a modern database against this archive and expect differences like these.

The source also records dates, not times. For 125 releases the auction and the announcement fell on the same day, and the data do not show their order within it.

What the original study found

The authors report that the market's central forecasts were very similar to, but more accurate than, the survey consensus (p. 1), and that, pooled across auctions, outcomes occurred about as often as prices implied, with no statistically significant miscalibration in their tests (pp. 23–30). Those are their results. We have not reproduced them, and the examples above can neither confirm nor contradict them. Our prespecified evaluation protocol sets out how a full-sample check should treat boundaries, open tails, totals that miss 100 and dependence between nearby releases, before anyone computes a headline number.

Use, cite and correct

Three distinct objects carry the authors' names: the original data and code package (readme dated 8 November 2005); Federal Reserve Bank of San Francisco Working Paper 2005-26, September 2005, whose pages we cite; and NBER Working Paper 11929, January 2006. Cite the one you used. Cite this edition as: Vultax Research, What the market expected: 153 historical economic-event auctions, Tallinn, 25 October 2026, https://vultax.com/research/economic-derivatives-auction-archive. The record page at https://vultax.com/archive/economic-derivatives lists the data files with their checksums. A DOI will be added here once the deposit is registered.

Charts exported from Figure 1 carry the authors' credit, record ID, units, edition and reveal state, so a slide keeps its source. Corrections are logged by ID in the version history, never silently overwritten; the correction path explains what to send.

Auction-only edition · JSON · Bins · CSV · Releases · CSV · Bins · JSON lines · Data dictionary · Citation file · BibTeX

Methods and data descriptor · Teaching exercise · Evaluation protocol · Research notes · Edition scope

Thank you

Refet Gürkaynak and Justin Wolfers did the original research and published their data and code alongside it. When we asked whether Vultax could republish the auction portion with credit, Justin Wolfers said yes. The survey and financial-market columns in their package come from other providers, so they are not here. His research-data page is worth exploring in its own right.

Vultax's part is the edition itself: we separated the 153 auction records from the original file, traced and checked every value against its source row, and built this edition, its explorer and the documentation that lets others verify the work.

ABOUT THE DATA

Auction data. Probabilities appear exactly as published, in percent, and are not rescaled. The lowest and highest ranges are open-ended; the source marks them with −9999 and 9999, which are not values. When an announced value falls exactly on a range boundary, we mark the boundary instead of assigning it to either side. Combined figures such as the 3.03 percent in the two lowest payroll ranges are sums of published range probabilities. Announced values are as recorded in the source, not later revisions. Survey and financial-market columns are not included.

Revisions. Current figures for December 2003 and December 2004 are month-on-month differences in the seasonally adjusted total nonfarm employment level, BLS series CES0000000001, retrieved from the BLS Public Data API on 22 October 2026 and rounded to the nearest thousand. They are context, not part of the edition.

SOURCES

- Refet S. Gürkaynak and Justin Wolfers, "Macroeconomic Derivatives: An Initial Analysis of Market-Based Macro Forecasts, Uncertainty, and Risk," Federal Reserve Bank of San Francisco Working Paper 2005-26, September 2005. Page numbers in this article refer to this version: https://www.frbsf.org/wp-content/uploads/wp05-26bk.pdf
- The same title as NBER Working Paper 11929, January 2006, doi:10.3386/w11929: https://www.nber.org/papers/w11929
- Justin Wolfers, research data, economic derivatives data and code (package dated 2005): https://users.nber.org/~jwolfers/data.php#economicderivativesdata
- US Bureau of Labor Statistics, Current Employment Statistics, series CES0000000001: https://data.bls.gov/timeseries/CES0000000001
- Selected arithmetic checked in Lean 4.35.0 — Selected exact-rational calculations and displayed-value consistency checks, checked by Lean 4.35.0, Lean kernel replay and Nanoda. Not a proof of the underlying observations, statistical conclusions, legal claims or production code.

CREDITS

- **Original research and auction data:** Refet S. Gürkaynak and Justin Wolfers. Permission to republish the auction data: Justin Wolfers, 21 October 2026.
- **Edition, writing, data checks and design:** Vultax Research, operated by OmniOS OÜ, Tallinn.

AI assistance was used for this edition's writing, code and data checks. Original research and data remain credited to their creators.

REFERENCES

Original economic-derivatives data and code https://users.nber.org/~jwolfers/data.php#economicderivativesdata

Published interactive research edition https://vultax.com/research/editions/economic-derivatives/index.html

Selected arithmetic checked in Lean 4.35.0 https://vultax.com/datasets/research-proof-library-2026-10-25/articles/economic-derivatives-auction-archive.md

Selected exact-rational calculations and displayed-value consistency checks, checked by Lean 4.35.0, Lean kernel replay and Nanoda. Not a proof of the underlying observations, statistical conclusions, legal claims or production code.

Per-article scope records and Lean package https://vultax.com/datasets/research-proof-library-2026-10-25/README.md

Dated per-article scope records, Lean source, frozen permitted public inputs, checker logs and a reproducible ZIP. Covers only the scope stated in each record, not the whole study.