



Asymmetric Okun Models and Unemployment Forecast Risk

Sity Komalla^{1*}, Alfa Rohmatin², Ahsen Yılmaz³

^{1,2} STIT Al Azami Cianjur, Indonesia

³ Ondokuz Mayıs University, Türkiye

Email: sitykomalla96@gmail.com¹

Article Info :

Received:
18-09-2026
Revised:
13-10-2026
Accepted:
16-10-2026

Abstract

This study evaluates whether asymmetric output growth improves one-quarter-ahead unemployment forecasts and whether any gains persist during major economic disruptions. Quarterly United States data spanning 1948–2024 were analyzed using autoregressive, linear Okun, and asymmetric Okun specifications. Predictive performance was assessed through 160 recursive pseudo-out-of-sample forecasts for 1985–2024, using root mean squared error and mean absolute error, alongside pandemic-related sensitivity analyses. All forecasts used revised historical data, not real-time information sets. Before 2020, the asymmetric model achieved an RMSE of 0.1779 percentage points, outperforming the linear model (0.1886) and autoregressive benchmark (0.1897). Across the full evaluation period, however, asymmetric RMSE increased to 1.1769, exceeding the linear model's 1.0502. Excluding 2020–2021 forecast targets restored its advantage over the linear model but yielded virtually identical accuracy to the autoregressive benchmark. Retrospectively omitting pandemic observations from estimation improved post-2021 performance, indicating substantial sensitivity to extreme shocks. The results show that statistically detectable output–unemployment asymmetry does not ensure robust predictive improvement. Forecasting applications should retain parsimonious benchmarks and evaluate model stability across contrasting macroeconomic regimes.

Keywords: unemployment forecasting; Okun's law; output contractions; pandemic; model stability.



©2022 Authors.. This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.
(<https://creativecommons.org/licenses/by-nc/4.0/>)

INTRODUCTION

Fiscal authorities and central banks use the relationship between output and unemployment to translate economic projections into assessments of labor market conditions. The practical attraction is considerable: output growth can summarize changes in demand, while unemployment is directly relevant to household income and the demand for public support. Yet a relationship that describes historical observations need not improve a forecast. This distinction is particularly consequential when a small number of extreme observations determine a model's estimated response to contractions. Ball, Leigh, and Loungani (2017) establish the empirical relevance of Okun's law, but its usefulness as a forecasting restriction remains a separate question.

This paper asks whether allowing different slopes for positive and negative output growth improves one-quarter-ahead predictions of changes in the United States unemployment rate. We examine 1948–2024, distinguish contemporaneous associations from lagged predictive equations, and compare three deliberately small models. The evaluation begins in 1985 and contains 160 forecast targets. Every coefficient estimate uses only earlier observations, although all observations come from the same revised data vintage. The exercise therefore measures pseudo-out-of-sample predictive performance rather than the performance of a live forecasting service.

Nonlinearity itself is not a new contribution. Research has examined asymmetric output–unemployment relationships across countries, business cycles, and historical periods (Bođ'a & Považanová, 2021; Donayre, 2022; Valadkhani & Smyth, 2015). Recent evidence covering 92 countries over 1980–2023 also documents substantial regime dependence (Sovbetov, 2025). We instead focus on the gap between a statistically detectable asymmetric association and the incremental accuracy of a deployable lagged equation. An autoregressive benchmark is essential to this comparison because unemployment changes may already contain much of the useful information attributed to output growth.

The findings are mixed in an economically informative way. Before 2020, the asymmetric model's root mean squared error is 0.1779 percentage points, compared with 0.1886 for the linear output model and 0.1897 for the autoregressive model. Across the entire evaluation period, the corresponding errors are 1.1769, 1.0502, and 1.1392. Removing 2020–2021 from the scoring period restores the asymmetric model's advantage over the linear model, but not over the autoregressive benchmark. These comparisons indicate that the case for adding asymmetry depends on the alternative model and the period being evaluated.

The analysis also separates two ways in which a crisis can alter forecast evaluation. Extreme outcomes create large forecast errors when they occur, and they can change estimated coefficients used afterward. Removing pandemic targets from an accuracy calculation addresses only the first channel. In a retrospective sensitivity exercise, removing pandemic observations from subsequent estimation substantially reduces post-2021 errors. That result is consistent with concerns about extreme observations in macroeconomic estimation, but it does not establish that deleting observations is an adequate method of modeling predictive uncertainty (Lenza & Primiceri, 2022).

The policy contribution is an assessment of a transparent forecasting rule, rather than an estimate of the effect of a particular intervention. Recent work in this journal on leading indicators illustrates the relevance of forecast evaluation to policy monitoring (Huang et al., 2026). Here, the evidence supports retaining a simple benchmark and reporting performance across crisis regimes before using an asymmetric output rule in labor market scenarios. It does not identify the employment multiplier of fiscal spending, the effect of interest rates, or the appropriate size of a stabilization program.

Okun's law summarizes a reduced-form relationship between output movements and unemployment. Output can adjust through hours, employment, labor force participation, and productivity, so a coefficient on output growth is not a technological constant. Gordon (2010) emphasizes productivity innovations in interpreting this relationship. Differences in labor market institutions and economic structure also limit the transferability of coefficients across countries: Ball et al. (2019) find marked cross-country heterogeneity, while Cazes et al. (2013) examine divergent unemployment responses to the global financial crisis. We consequently restrict the empirical claim to a single national labor market.

An asymmetric response is economically plausible because a contraction may induce rapid job destruction, whereas renewed production can initially be accommodated through hours and productivity. The jobless recovery literature examines several versions of this adjustment problem (Elroukh et al., 2020; Panovska, 2017). Burda and Hunt (2011) further show why one country's employment response during a major downturn need not resemble another's. None of these mechanisms is separately identified by our two-series specification. They motivate allowing the data to depart from a common slope, without licensing a structural interpretation of the estimated difference.

The empirical literature offers several ways to model that departure. Valadkhani and Smyth (2015) study switching and asymmetric behavior in the United States over 1948–2015. Nebot et al. (2019) examine nonlinearity, and Donayre (2022) studies behavior across business cycles. Bod'a and Považanová (2021) extend the asymmetry question across Organisation for Economic Co-operation and Development (OECD) economies; Mussida and Zanin (2023) investigate both asymmetry and instability in European countries. Our zero-growth split is simpler than a model with estimated transition probabilities or an unknown threshold. Its advantage is that the additional restriction can be evaluated with one extra slope and a clearly specified benchmark.

Timing matters independently of functional form. Delayed unemployment adjustment is examined by Obst (2022), while Grant (2018) and Owyang and Sekhposyan (2012) study the relationship around the Great Recession. The forecasting equations below therefore include the preceding unemployment change in every specification. This prevents a comparison in which the nonlinear model benefits simply from additional unemployment dynamics. It also makes clear that a slope on lagged output growth answers a different question from a slope on output observed during the same quarter as the unemployment outcome.

Measurement choices can generate additional variation in reported coefficients. Bod'a and Považanová (2023) demonstrate the importance of assessing the credibility of gap-based estimates. Hamilton (2018) identifies problems with a commonly used method of extracting trends. We work with changes and growth rates rather than estimated unemployment and output gaps. This avoids choosing potential output or a smoothing parameter, although it does not remove revisions, measurement error,

or structural change. Porras-Arena and Martín-Román (2023) and Sovbetov (2025) provide further reasons to avoid interpreting one coefficient as a universally applicable rule.

The pandemic requires a separate evaluation because abrupt changes in production and employment need not propagate like ordinary cyclical fluctuations. Gupta et al. (2023) analyze labor market outcomes in relation to social distancing policies, while Lenza and Primiceri (2022) distinguish coefficient estimation from the uncertainty surrounding future outcomes after extreme observations. More generally, Mathy (2024) revisits the interpretation of slow recoveries. Our contribution is narrower: we report what happens to a small forecasting rule when the pandemic enters both its evaluation period and its estimation history. We do not attribute the resulting instability to one policy or one labor market mechanism.

RESEARCH METHODS

Sources and sample construction

The analysis uses real gross domestic product (GDP) from the United States Bureau of Economic Analysis (BEA) and the seasonally adjusted unemployment rate from the United States Bureau of Labor Statistics (BLS), distributed through Federal Reserve Economic Data (FRED). Real gross domestic product is series GDPC1, measured in billions of chained 2017 dollars at a seasonally adjusted annual rate. The unemployment rate is series UNRATE, measured monthly in percent. The retrieval date is September 11, 2026. The source tables identify updates of August 26, 2026 for GDPC1 and September 4, 2026 for UNRATE (BEA, 2026; BLS, 2026).

The main sample covers 1948Q1–2024Q4. It contains 924 monthly unemployment observations and 308 quarterly unemployment averages. We retain the arithmetic average of the three months in each quarter; all quarters in this interval are complete. Real output observations from 1947 provide the presample level needed to construct initial growth. No interpolation, winsorization, seasonal adjustment, or gap filtering is performed. Ending in 2024 supplies a balanced sequence of complete calendar years; the retrieved unemployment series contains an unavailable observation for October 2025, which is not imputed or used.

Quarterly output growth is 100 times the difference in the natural logarithm of real output. It is not annualized, despite the annual-rate convention used for the underlying output level. Unemployment changes are measured in percentage points. The first unemployment difference is available in 1948Q2, and its first lag becomes available in 1948Q3. All regressions use the common sample beginning in 1948Q3, with 306 observations through 2024Q4. Table 1 documents the variables and their construction; the supplementary material provides a detailed audit trail and historical summary statistics.

Table 1 Variables And Analysis Sample

Variable	Definition	Unit or coverage
Real output	BEA series GDPC1	Quarterly; chained 2017 USD billions; annual rate
Unemployment	BLS series UNRATE; mean of three months	Percent; 924 monthly values
Output growth	$100 \times$ quarterly change in log output	Percent; not annualized
Unemployment change	Current minus preceding quarterly rate	Percentage points
Positive growth	Maximum of growth and zero	Percent
Negative growth	Minimum of growth and zero	Percent; negative sign retained
Observation interval	1948Q1–2024Q4	308 quarters
Regression interval	1948Q3–2024Q4	306 complete rows
Forecast targets	1985Q1–2024Q4	160 quarters

Source: calculations from BEA GDPC1 and BLS UNRATE.

Figure 1 places the two variables on their original economic scales. Across 1948–2024, the unemployment rate averages 5.6832%, and quarterly output growth averages 0.7693%. There are 42

quarters of negative output growth. The 2020–2021 block contains only eight quarters, but the standard deviation of quarterly unemployment changes within that block is 3.9234 percentage points, compared with 0.2068 in 1984–2007. These comparisons establish that a long sample does not necessarily distribute statistical influence evenly over time. They do not imply that observations should be removed from the baseline analysis.

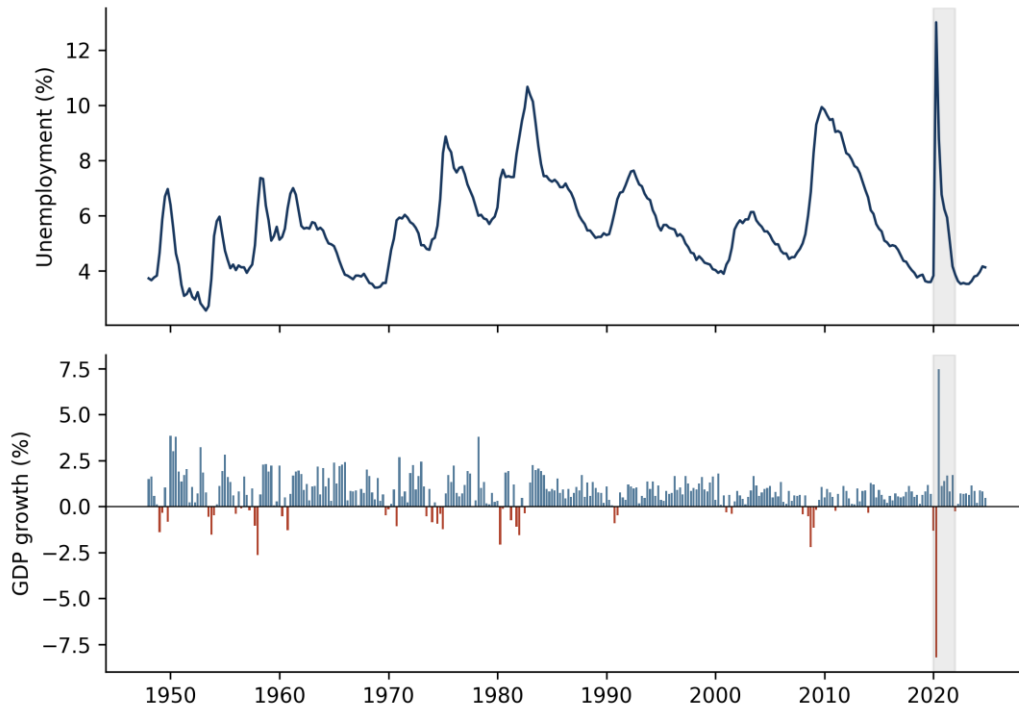


Figure 1 Observed Unemployment and Output Growth

Source: calculations from BEA GDPC1 and BLS UNRATE

Contemporaneous and predictive specifications

Let u denote the quarterly unemployment rate and g denote quarterly output growth. Define positive growth as the maximum of g and zero, and negative growth as the minimum of g and zero. The negative component retains its negative sign. The two components sum exactly to observed growth, and their use imposes continuity at zero. Zero is specified as the boundary between output expansion and contraction; it is not estimated by searching for the most significant threshold. A negative-growth quarter is not equated with an officially dated recession.

$$g_t = 100[\ln(Y_t) - \ln(Y_{t-1})] \quad (1)$$

The contemporaneous specification relates the unemployment change to its first lag and the two components of current output growth. We use this equation only to describe associations. Output during the target quarter is not available at the beginning of that quarter, so its fitted values cannot be treated as one-quarter-ahead forecasts. The common-slope null is equality of the positive- and negative-growth coefficients. Testing that restriction addresses asymmetry within this particular model; it does not test all possible forms of nonlinearity.

$$\Delta u_t = \alpha + \phi \Delta u_{t-1} + \beta^+ \max(g_t, 0) + \beta^- \min(g_t, 0) + \varepsilon_t \quad (2)$$

Three predictive equations are then estimated. The autoregressive (AR) model includes an intercept and the previous unemployment change. The linear output model adds the previous quarter's output growth. The asymmetric model replaces that single growth term with its positive and negative components. Thus the linear model is nested within the asymmetric model through a slope-equality restriction, while the autoregressive model provides a parsimonious reference that does not use output

at all. The asymmetric equation is shown below; imposing equal output slopes gives the linear model, and setting both to zero gives the autoregressive model.

$$\Delta u_t = \alpha + \varphi \Delta u_{t-1} + \beta^+ \max(g_{t-1}, 0) + \beta^- \min(g_{t-1}, 0) + \varepsilon_t \quad (3)$$

Coefficients are estimated by ordinary least squares. For reported full-sample and subsample equations, standard errors use a Bartlett heteroskedasticity and autocorrelation consistent (HAC) covariance estimator with four quarterly lags and a degrees-of-freedom adjustment (Newey & West, 1987). Calendar spacing is preserved when pandemic observations are excluded: the final quarter of 2019 is not treated as adjacent to the first quarter of 2022. We report two-sided normal-approximation tests of slope equality and examine bandwidths of zero and eight in the supplement. These are conditional linear-projection tests, not causal identification tests.

Recursive evaluation and sensitivity analysis

Forecast evaluation covers 1985Q1–2024Q4. At each target date, the initial estimation sample is extended through the previous quarter, the model is re-estimated, and a prediction of the target-quarter unemployment change is recorded. The initial window contains 146 usable observations. We do not optimize the lag order or the growth threshold against forecast results. The starting date supplies a substantial training history and places evaluation after the early-1980s change in United States output volatility discussed by McConnell and Perez-Quiros (2000). It is a design choice, not an estimated break date.

All variables are measured in a common revised vintage. In a real-time implementation, both release dates and revisions would matter, and the previous quarter's initial output estimate would become available only after that quarter ended. Our use of lagged predictors prevents direct use of target outcomes in estimation, but it does not recreate the information set available at every historical forecast origin. The resulting forecasts should be interpreted as a controlled comparison of predictive equations under a common retrospective information set. Their reported accuracy is not a claim about contemporaneous operational performance.

We report root mean squared error (RMSE), mean absolute error (MAE), and mean signed error. Forecast error is the actual unemployment change minus its prediction, so a negative mean error indicates average overprediction. Squared loss emphasizes unusually large mistakes, while absolute loss describes a less tail-sensitive criterion. The distinction follows the general logic of forecast comparison developed by Diebold and Mariano (1995). Because these models are nested and repeatedly estimated, conventional comparisons require care; Clark and West (2007) explain why unadjusted mean squared error comparisons need not supply a valid nested-model significance test.

Our uncertainty assessment is therefore presented as a descriptive paired block-bootstrap interval for realized loss differences, not as a formal nested-model selection test. For each comparison, we resample eight-quarter blocks of paired forecast losses 4,999 times and report percentile intervals. Blocks cannot cross a removed calendar interval. The procedure follows the dependence-preserving principle of Künsch (1989), but does not re-estimate the predictive models within each bootstrap draw. It cannot establish exact coverage under a crisis-induced distributional break. All point rankings are reported whether or not the interval excludes zero.

We show the full evaluation, the pre-2020 interval, pandemic targets in 2020–2021, targets outside that block, and the 2022–2024 period. Two additional estimation schemes use a rolling 80-quarter window and an expanding window that retrospectively omits training rows from 2020Q1 through 2022Q1. The latter also removes the first row whose lagged variables originate in 2021Q4. It is an ex post diagnostic, not a rule assumed to have been known before the pandemic. Neither sensitivity changes the realized outcome or suppresses pandemic forecast targets.

RESULTS AND DISCUSSION

Asymmetric fit and predictive slopes

The contemporaneous output–unemployment relationship is asymmetric before the pandemic and in the full sample. Table 2 shows a pre-2020 positive-growth slope of -0.1670 and a negative-growth slope of -0.3653 . Holding the lagged unemployment change fixed, a one-percentage-point decrease in quarterly growth within the contraction region is associated with a 0.3653-percentage-point

increase in the unemployment change. The corresponding difference between slopes is -0.1983 , with a standard error of 0.0557 . The equality restriction is rejected with $p = 0.0004$.

Table 2 Contemporaneous Output Association

Parameter	1948–2024	1948–2019	Excluding 2020–2021
Intercept	0.1151 (0.0435)	0.1087 (0.0263)	0.1069 (0.0253)
Lagged unemployment change	-0.0053 (0.1110)	0.4545 (0.0379)	0.4589 (0.0367)
Positive output growth	-0.2544 (0.0558)	-0.1670 (0.0250)	-0.1641 (0.0246)
Negative output growth	-0.8725 (0.1564)	-0.3653 (0.0431)	-0.3648 (0.0429)
Negative minus positive slope	-0.6181	-0.1983	-0.2007
Slope equality p-value	0.0000	0.0004	0.0003
R-squared	0.7121	0.6788	0.6761
Observations	306	286	298

Source: calculations from BEA GDPC1 and BLS UNRATE

Including 2020–2024 makes the contemporaneous contraction slope much larger in magnitude, at -0.8725 , while the positive-growth slope is -0.2544 . The model's R-squared is 0.7121 . Figure 2 shows why the expanded range of observations matters: the pandemic combines unusually large movements in both series. The plotted lines hold the lagged unemployment change at zero and are not unconditional fitted regressions through the scatter. A stronger full-sample fit is compatible with a less reliable forecasting equation, since current output and lagged output contain different information.

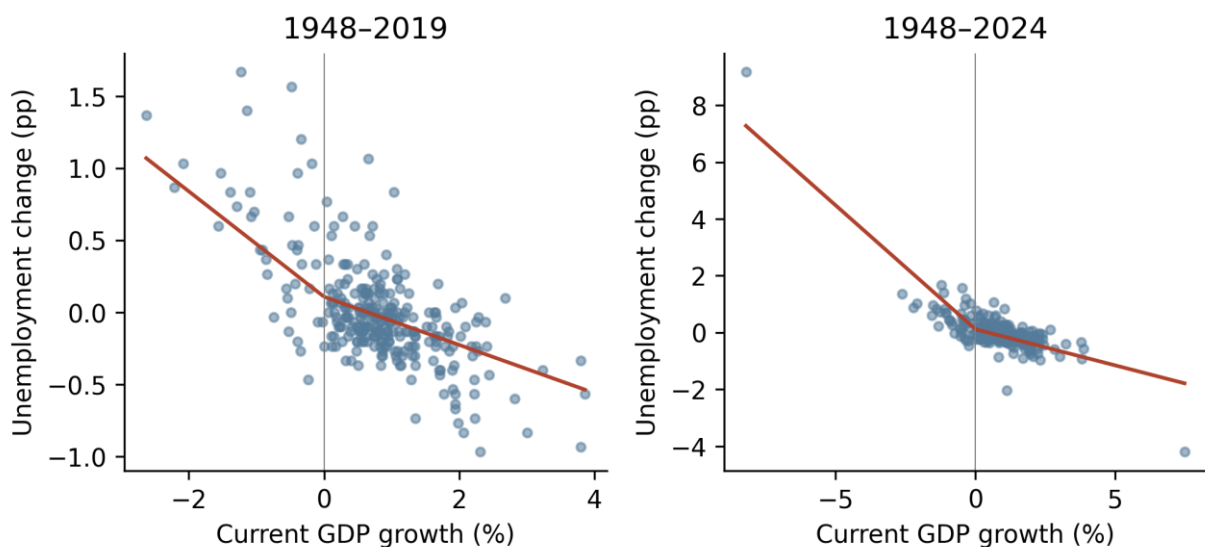


Figure 2 Contemporaneous Output Association Before And Including The Pandemic

Table 3 makes that distinction explicit. In the pre-2020 predictive equation, the positive lagged-growth slope is -0.0744 and the negative slope is -0.2766 . Their difference is -0.2022 , with $p = 0.0349$. The lagged unemployment-change coefficient is 0.4181 . When the full sample is used, that persistence coefficient becomes -0.2632 , the positive-growth slope becomes -0.3659 , and the negative-growth slope becomes -0.1081 . The slope difference changes sign and is not statistically distinguishable from zero at conventional levels, with $p = 0.1658$.

Table 3 Lagged Asymmetric Predictive Equation

Parameter	1948–2024	1948–2019	Excluding 2020–2021
Intercept	0.3166 (0.1388)	0.0367 (0.0306)	0.0368 (0.0294)
Lagged unemployment change	-0.2632 (0.2956)	0.4181 (0.0695)	0.4202 (0.0663)
Positive output growth lag	-0.3659 (0.1460)	-0.0744 (0.0301)	-0.0725 (0.0293)
Negative output growth lag	-0.1081 (0.2081)	-0.2766 (0.0927)	-0.2738 (0.0913)
Negative minus positive slope	0.2579	-0.2022	-0.2013
Slope equality p-value	0.1658	0.0349	0.0337
R-squared	0.1363	0.4797	0.4787
Observations	306	286	298

Source: calculations from BEA GDPC1 and BLS UNRATE.

Excluding 2020–2021 target observations from coefficient estimation produces slopes close to the pre-2020 values. The positive and negative coefficients are -0.0725 and -0.2738 , respectively, and the lagged-change coefficient is 0.4202 . This exercise concerns estimation rather than forecast scoring. It shows that the full-sample sign reversal is highly sensitive to a small calendar block, but does not determine whether the change represents altered behavior, an exceptional shock sequence, measurement effects, or omitted variables. We consequently avoid describing the full-sample slopes as evidence of a stable new regime.

Forecast Accuracy Depends on The Evaluation Period

Before 2020, the asymmetric equation reduces RMSE by 5.6% relative to the linear output equation and by 6.2% relative to the autoregressive benchmark. MAE is also lower: 0.1409 compared with 0.1502 and 0.1476 . The gain is not simply a consequence of comparing models with different unemployment lags, because the same lagged unemployment change appears in all three. Table 4 reports these comparisons alongside the full and pandemic-period results. The relevant conclusion is an incremental pre-pandemic gain within this limited set of models.

Table 4 Recursive Expanding-Window Forecast Accuracy

Targets	Model	N	RMSE	MAE	Mean error
1985–2024	AR	160	1.1392	0.2933	-0.0516
1985–2024	Linear	160	1.0502	0.2951	-0.0779
1985–2024	Asymmetric	160	1.1769	0.2961	-0.0786
1985–2019	AR	140	0.1897	0.1476	-0.0130
1985–2019	Linear	140	0.1886	0.1502	-0.0434
1985–2019	Asymmetric	140	0.1779	0.1409	-0.0287
Outside 2020–2021	AR	152	0.1864	0.1448	-0.0122
Outside 2020–2021	Linear	152	0.1986	0.1546	-0.0449
Outside 2020–2021	Asymmetric	152	0.1868	0.1447	-0.0330
2020–2021	AR	8	5.0295	3.1133	-0.8015
2020–2021	Linear	8	4.6162	2.9656	-0.7044
2020–2021	Asymmetric	8	5.1998	3.1726	-0.9443
2022–2024	AR	12	0.1434	0.1127	-0.0024
2022–2024	Linear	12	0.2914	0.2059	-0.0622
2022–2024	Asymmetric	12	0.2694	0.1884	-0.0837

Source: calculations from BEA GDPC1 and BLS UNRATE

The ranking reverses over the full 160-quarter evaluation. The asymmetric equation's RMSE is 1.1769, which is 12.1% above the linear model's 1.0502. Absolute errors are much closer: MAE is 0.2961 for the asymmetric model and 0.2951 for the linear model. This difference between loss functions indicates that a few exceptionally large errors contribute substantially to the RMSE ranking. Figure 3 contrasts full-period performance with the evaluation that excludes pandemic targets. It also shows why the autoregressive benchmark changes the interpretation of the nonlinear model's apparent gain.

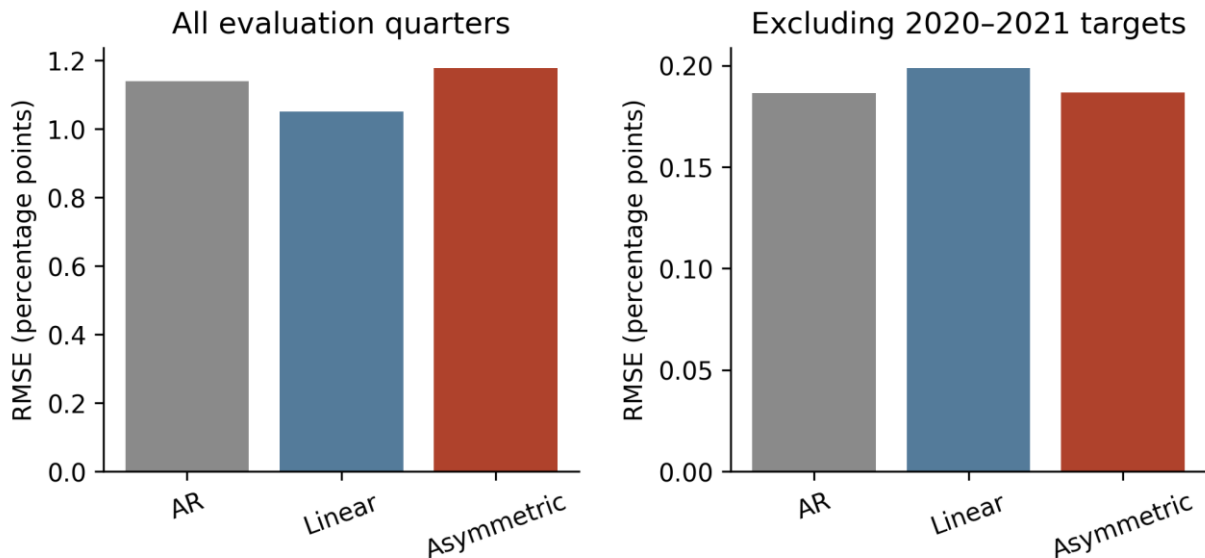


Figure 3 Forecast RMSE With And Without Pandemic Targets

For the 152 targets outside 2020–2021, the asymmetric equation has RMSE of 0.1868, compared with 0.1986 for the linear equation. However, the autoregressive model has RMSE of 0.1864. Its MAE of 0.1448 is also almost identical to the asymmetric model's 0.1447. Thus an improvement over a linear output equation does not establish an improvement over a model that omits output. This is a practical restriction on the claim that additional nonlinear output terms make an unemployment forecasting system more informative.

The paired loss intervals in Table 5 reinforce the need to specify the benchmark. Before 2020, the squared-loss gain relative to the linear equation is 0.00390, with a descriptive interval of [0.00126, 0.00777]. The corresponding interval relative to the autoregressive model includes zero. Outside pandemic targets, the gain relative to the linear equation remains positive, while the comparison with the autoregressive model is centered near zero. Full-period intervals are wide and span zero. We do not interpret these intervals as a uniformly valid rejection of equal predictive ability under structural change.

Table 5 Squared-Loss Gains Of The Asymmetric Model

Targets	Benchmark	Gain	95% interval
1985–2024	AR	-0.08722	[-0.30529, 0.03913]
1985–2024	Linear	-0.28210	[-0.87862, 0.02216]
1985–2019	AR	0.00432	[-0.00297, 0.01405]
1985–2019	Linear	0.00390	[0.00126, 0.00777]
Outside 2020–2021	AR	-0.00013	[-0.00617, 0.01312]
Outside 2020–2021	Linear	0.00457	[0.00148, 0.00770]

Source: calculations from BEA GDPC1 and BLS UNRATE

Figure 4 reports cumulative gains relative to the linear output model. Positive values favor the asymmetric equation. The squared-error series rises gradually before 2020 and then falls sharply during the pandemic, whereas the cumulative absolute-error comparison is less dominated by the largest observations. This display makes the timing of the ranking reversal visible without selecting a favorable

final average. Both paths use every target in the evaluation interval; no forecast error is trimmed or capped.

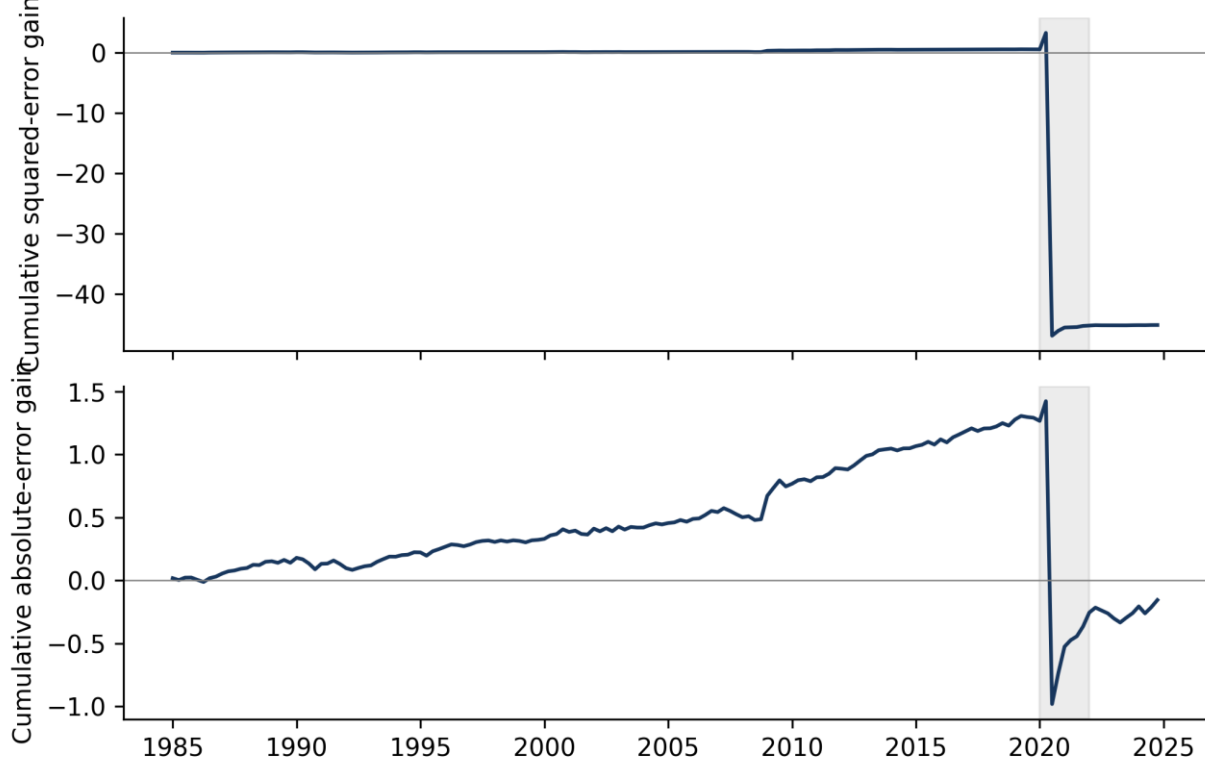


Figure 4 Cumulative forecast loss gains relative to the linear model

Positive values favor the asymmetric model. Shading identifies 2020–2021. Upper panel is in squared percentage points; lower panel is in percentage points.

Source: calculations from BEA GDPC1 and BLS UNRATE

Pandemic Extrapolation and Subsequent Estimation

The largest errors are not confined to the initial unemployment increase. In 2020Q2, the actual quarterly unemployment change is 9.1667 percentage points. The asymmetric equation predicts an increase of 0.5006, compared with 0.3443 for the linear equation and 0.1525 for the autoregressive model. The nonlinear term helps slightly in that quarter, but all three miss most of the increase. In 2020Q3, unemployment falls by 4.2000 points, whereas the asymmetric equation predicts an increase of 7.4392. Its absolute error is then 11.6392 points, larger than those of both benchmarks.

This sequence illustrates the danger of extrapolating a lagged cyclical relationship through an abrupt shutdown and reversal. A very large increase in unemployment and a severe output contraction enter the next forecast as predictors of continued deterioration. The result does not contradict a negative contemporaneous association between output and unemployment. It shows that persistence inferred from ordinary episodes can generate a severely incorrect prediction when the timing and scale of the episode differ. Supplementary Table S4 reports every pandemic forecast so that this interpretation is not based on selectively quoting one error.

Post-pandemic performance also depends on how the pandemic enters estimation. For 2022–2024, the standard expanding asymmetric equation has RMSE of 0.2694, compared with 0.1434 for the autoregressive model. Retrospectively excluding pandemic training rows reduces these values to 0.1210 and 0.1213, respectively. The linear model's RMSE similarly falls from 0.2914 to 0.1224. Table 6 reports the alternative estimation schemes. These are substantial changes, but the comparison contains only 12 post-pandemic targets and cannot establish durable superiority.

Table 6 RMSE Under Alternative Training Schemes

Training scheme	Model	1985–2019	1985–2024	2022–2024
Expanding	AR	0.1897	1.1392	0.1434
Expanding	Linear	0.1886	1.0502	0.2914
Expanding	Asymmetric	0.1779	1.1769	0.2694
Rolling 80	AR	0.1931	1.3405	0.1928
Rolling 80	Linear	0.1858	1.2382	0.5731
Rolling 80	Asymmetric	0.1745	1.6856	0.5369
Pandemic training excluded	AR	0.1897	1.0937	0.1213
Pandemic training excluded	Linear	0.1886	1.0425	0.1224
Pandemic training excluded	Asymmetric	0.1779	1.0826	0.1210

Source: calculations from BEA GPC1 and BLS UNRATE

A rolling window does not automatically solve the problem. The rolling asymmetric model improves on the rolling linear model before 2020, with RMSE of 0.1745 versus 0.1858. Across all targets, however, its RMSE increases to 1.6856. Over 2022–2024 it remains 0.5369, well above the rolling autoregressive model's 0.1928. Recent extreme observations occupy a larger share of an 80-quarter sample than of an expanding sample. The rolling-coefficient figure in the supplement documents the resulting instability without treating each window as an independent statistical test.

Implications For Policy Forecasting

The first implication is that model assessment should distinguish explanatory fit, incremental forecasting accuracy, and operational usefulness. Table 2 supports an asymmetric contemporaneous association. Tables 4 and 5 provide qualified evidence of an incremental pre-pandemic forecasting gain. Neither finding establishes that the model should control a policy decision. A forecasting rule can remain informative as one input to a labor market assessment while failing to outperform a simple benchmark under the losses and episodes that matter most to its user.

The second implication concerns benchmark selection. Comparing a nonlinear output equation only with a linear output equation can make nonlinear modeling appear more valuable than it is. Outside pandemic targets, the autoregressive model nearly matches the asymmetric model while using fewer inputs and parameters. For an agency with limited analytical capacity, this is relevant to the cost of maintaining additional data series and model variants. We do not estimate those costs; the empirical result is the absence of a meaningful point-accuracy improvement over the autoregressive reference in that scoring interval.

The third implication is that crisis treatment should be explicit in both estimation and evaluation. Removing crisis errors from an accuracy report can restore a favorable ranking while leaving crisis observations embedded in later coefficient estimates. Conversely, excluding extreme estimation observations can improve point forecasts while producing an inadequate picture of future uncertainty. The latter distinction is central to Lenza and Primiceri (2022). Our retrospective deletion exercise identifies sensitivity and motivates further work with volatility-aware methods; it is not a recommendation to erase exceptional events from institutional forecasting records.

The fourth implication concerns communication. A policy brief should state the target, horizon, predictor timing, data vintage, benchmark, loss function, and treatment of exceptional periods. An output contraction in this paper means a negative quarterly growth observation, not an officially classified recession or an identified negative demand shock. The estimated slopes cannot be used to calculate how much fiscal expenditure would reduce unemployment. Any application to another country would additionally require attention to informality, participation, measurement conventions, and labor market structure, as emphasized by cross-country research (Ball et al., 2019).

CONCLUSION

The findings show that asymmetric output dynamics provide a meaningful explanation of unemployment movements in United States quarterly data from 1948 to 2024 and offer modest forecasting improvements over a linear output specification in the pre-2020 period. However, this advantage disappears when the pandemic period is included and becomes negligible compared with a

parsimonious autoregressive benchmark outside crisis-related observations. These results indicate that strong historical nonlinear relationships do not necessarily translate into stable forecasting gains, particularly when economic conditions are affected by exceptional structural shocks.

Accordingly, forecasting models based on output–unemployment relationships should be evaluated against simple benchmarks and tested for sensitivity to extreme observations, structural change, and alternative estimation periods. The evidence supports the use of nonlinear specifications as complementary forecasting tools rather than universally superior models. Reliable policy forecasting therefore requires attention not only to statistical fit but also to forecast robustness, model stability, and performance across substantially different macroeconomic environments.

REFERENCES

- Ball, L., Furceri, D., Leigh, D. & Loungani, P. (2019) Does one law fit all? Cross-country evidence on Okun's law. *Open Economies Review*, 30, 841–874. <https://doi.org/10.1007/s11079-019-09549-3>
- Ball, L., Leigh, D. & Loungani, P. (2017) Okun's law: fit at 50? *Journal of Money, Credit and Banking*, 49(7), 1413–1441. <https://doi.org/10.1111/jmcb.12420>
- Bod'a, M. & Považanová, M. (2021) Output-unemployment asymmetry in Okun coefficients for OECD countries. *Economic Analysis and Policy*, 69, 307–323. <https://doi.org/10.1016/j.eap.2020.12.004>
- Bod'a, M. & Považanová, M. (2023) How credible are Okun coefficients? The gap version of Okun's law for G7 economies. *Economic Change and Restructuring*, 56, 1467–1514. <https://doi.org/10.1007/s10644-022-09438-9>
- Burda, M.C. & Hunt, J. (2011) What explains the German labor market miracle in the Great Recession? *Brookings Papers on Economic Activity*, 2011(1), 273–319. <https://doi.org/10.1353/eca.2011.0004>
- Cazes, S., Verick, S. & Al Hussami, F. (2013) Why did unemployment respond so differently to the global financial crisis across countries? Insights from Okun's law. *IZA Journal of Labor Policy*, 2, 10. <https://doi.org/10.1186/2193-9004-2-10>
- Clark, T.E. & West, K.D. (2007) Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1), 291–311. <https://doi.org/10.1016/j.jeconom.2006.05.023>
- Diebold, F.X. & Mariano, R.S. (1995) Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263. <https://doi.org/10.1080/07350015.1995.10524599>
- Donayre, L. (2022) On the behavior of Okun's law across business cycles. *Economic Modelling*, 112, 105858. <https://doi.org/10.1016/j.econmod.2022.105858>
- Elroukh, A.W., Nikolsko-Rzhevskyy, A. & Panovska, I. (2020) A look at jobless recoveries in G7 countries. *Journal of Macroeconomics*, 64, 103206. <https://doi.org/10.1016/j.jmacro.2020.103206>
- Gordon, R.J. (2010) Okun's law and productivity innovations. *American Economic Review*, 100(2), 11–15. <https://doi.org/10.1257/aer.100.2.11>
- Grant, A.L. (2018) The Great Recession and Okun's law. *Economic Modelling*, 69, 291–300. <https://doi.org/10.1016/j.econmod.2017.10.002>
- Gupta, S., Montenegro, L., Nguyen, T., Lozano-Rojas, F., Schmutte, I., Simon, K., Weinberg, B.A. & Wing, C. (2023) Effects of social distancing policy on labor market outcomes. *Contemporary Economic Policy*, 41(1), 166–193. <https://doi.org/10.1111/coep.12582>
- Hamilton, J.D. (2018) Why you should never use the Hodrick-Prescott filter. *The Review of Economics and Statistics*, 100(5), 831–843. https://doi.org/10.1162/rest_a_00706
- Huang, Y., Liu, Y. & Ding, M. (2026) Tracking Hong Kong economy with a novel leading index. *Contemporary Economic Policy*, advance online publication. <https://doi.org/10.1111/coep.70051>
- Künsch, H.R. (1989) The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17(3), 1217–1241. <https://doi.org/10.1214/aos/1176347265>
- Lenza, M. & Primiceri, G.E. (2022) How to estimate a vector autoregression after March 2020. *Journal of Applied Econometrics*, 37(4), 688–699. <https://doi.org/10.1002/jae.2895>

- Mathy, G. (2024) Are jobless recoveries history? Okun's law, insufficient stimulus, and slow recoveries. *Review of Keynesian Economics*, 12(4), 435–452. <https://doi.org/10.4337/roke.2024.0002>
- McConnell, M.M. & Perez-Quiros, G. (2000) Output fluctuations in the United States: what has changed since the early 1980's? *American Economic Review*, 90(5), 1464–1476. <https://doi.org/10.1257/aer.90.5.1464>
- Mussida, C. & Zanin, L. (2023) Asymmetry and (in-)stability of Okun's coefficients in nine European countries. *The Journal of Economic Asymmetries*, 28, e00313. <https://doi.org/10.1016/j.jeca.2023.e00313>
- Nebot, C., Beyaert, A. & García-Solanes, J. (2019) New insights into non-linearity of Okun's law. *Economic Modelling*, 82, 202–210. <https://doi.org/10.1016/j.econmod.2019.01.005>
- Newey, W.K. & West, K.D. (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708. <https://doi.org/10.2307/1913610>
- Obst, T. (2022) Dynamic version of Okun's law in the EU15 countries—The role of delays in the unemployment-output nexus. *Scottish Journal of Political Economy*, 69(2), 225–241. <https://doi.org/10.1111/sjpe.12305>
- Owyang, M.T. & Sekhposyan, T. (2012) Okun's law over the business cycle: was the Great Recession all that different? *Federal Reserve Bank of St. Louis Review*, 94(5), 399–418. https://fraser.stlouisfed.org/title/review-federal-reserve-bank-st-louis-820/september-october-2012-620781?page=87&toc_id=634619
- Panovska, I.B. (2017) What explains the recent jobless recoveries? *Macroeconomic Dynamics*, 21(3), 708–732. <https://doi.org/10.1017/S1365100515000656>
- Porrás-Arena, M.S. & Martín-Román, A.L. (2023) The correlation between unemployment and economic growth in Latin America—Okun's law estimates by country. *International Labour Review*, 162, 171–198. <https://doi.org/10.1111/ilr.12398>
- Sovbetov, I. (2025) The dynamics of Okun's law: cross-country analysis across economic cycles (1980–2023). *Empirica*, 52, 687–720. <https://doi.org/10.1007/s10663-025-09662-2>
- U.S. Bureau of Economic Analysis (BEA) (2026) Real gross domestic product [GDPC1]. FRED, Federal Reserve Bank of St. Louis. Retrieved September 11, 2026. <https://fred.stlouisfed.org/series/GDPC1>
- U.S. Bureau of Labor Statistics (BLS) (2026) Unemployment rate [UNRATE]. FRED, Federal Reserve Bank of St. Louis. Retrieved September 11, 2026. <https://fred.stlouisfed.org/series/UNRATE>
- Valadkhani, A. & Smyth, R. (2015) Switching and asymmetric behaviour of the Okun coefficient in the US: evidence for the 1948–2015 period. *Economic Modelling*, 50, 281–290. <https://doi.org/10.1016/j.econmod.2015.07.001>