Federal Reserve Bank of Cleveland Working Paper Series

Are Revisions to State-Level GDP Data in the US Well Behaved?

James Mitchell and Taylor Shiroff

Working Paper No. 25-11

April 2025

Suggested citation: Mitchell, James and Taylor Shiroff. 2025. "Are Revisions to State-Level GDP Data in the US Well Behaved?" Working Paper No. 25-11. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202511.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Working papers of the Federal Reserve Bank of Cleveland are preliminary materials circulated to stimulate discussion and critical comment on research in progress. They may not have been subject to the formal editorial review accorded official Federal Reserve Bank of Cleveland publications.

See more working papers at: www.clevelandfed.org/research. Subscribe to email alerts to be notified when a new working paper is posted at: https://www.clevelandfed.org/subscriptions.

This work is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International. To view a copy of this license, visit https://creativecommons.org/licenses/by-nc-nd/4.0/

Are Revisions to State-Level GDP Data in the US Well Behaved?*

James Mitchell† Taylor Shiroff‡

April 14, 2025

Abstract

No, first estimates of state GDP growth are not rational forecasts, except for Georgia. Revisions to first estimates of state-level GDP growth tend to be biased, large, and/or predictable using information known at the time of the first estimate.

Keywords: Data revisions; Real-time data; State GDP; Forecast efficiency

JEL Codes: E01, R11

1 Introduction

"Official" estimates of macroeconomic variables like GDP are continuously revised by the statistical office, reflecting, among other things, the arrival of additional information and reconciliation with alternative surveys. As Croushore (2011) reviews, data revisions can be both statistically and economically meaningful. Existing real-time data research has focused on analyzing the properties of revisions to macroeconomic data at the *national* level. This is despite growing interest in tracking and understanding the evolution of regional economies, including at the state level in the US; for example, see Hamilton and Owyang (2012), Arias

*The views expressed herein are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System. Thanks to Ed Knotek for helpful comments.

†Research Department, Federal Reserve Bank of Cleveland. Email address: james.mitchell@clev.frb.org

‡Research Department, Federal Reserve Bank of Cleveland. Email address: taylor.shiroff@clev.frb.org

et al. (2016), Bokun et al. (2023), and Baumeister et al. (2024). State GDP data are also used to help identify structural macroeconomic relationships, evidencing the need to understand the drivers of cross-state variation and the possible role of data revisions; for example, see Nakamura and Steinsson (2014).

This paper provides the first evaluation of the empirical properties of state-level GDP data in the US, as published by the Bureau of Economic Analysis (BEA). Extending Aruoba's (2008) study of US GDP revisions, and borrowing that paper's language and methodology, we test if revisions to first estimates of GDP growth in state i ($i = 1, ..., N$) in quarter t ($t = 1, ..., T$), rev_{it}, are "well behaved." Well behaved implies that the first estimates are rational forecasts, in the sense that they optimally incorporate all available information. Revisions are then classified as "news," rather than "noise" (cf. Mankiw and Shapiro (1986)). Following Aruoba (2008), we test three properties of rationality (under quadratic loss):

(P1) $E(rev_{it}) = 0$;

(P2) $SD(rev_{it})$ is small;

(P3) $E(rev_{it}|I_{t+1}) = 0$.

(P1) implies that revisions are mean zero – they are unbiased. (P2) says that, under rationality, we expect the standard deviation of revisions to be small relative to the variance of the "final" state GDP estimate. (P3) implies efficiency of the revisions process: Revisions to GDP growth estimates, if well behaved, should be unpredictable given information, I_{t+1}, known at the time of the first data release, typically made one quarter after the quarter to which the GDP data are related.

When data revisions satisfy (P1) through (P3), users of state-level GDP data can use first release data "as is," albeit they should acknowledge the data uncertainty, perhaps communicated via fan charts; cf. Galvão and Mitchell (2023). However, if data revisions are biased and/or predictable, then use of first estimates of state GDP biases inference. More-over, as Bokun et al. (2023) conclude, the use of real-time data then becomes essential when evaluating state-level forecasting models via out-of-sample experiments.

2 Real-time state-level GDP data

The BEA first published quarterly estimates of state-level real GDP in September 2015. These quarterly data supplemented existing annual data. Our real-time analysis focuses on the quarterly data, despite their shorter historical span, given the interest in tracking the US economy and its regions intra-year.

We analyze state-level GDP data vintages for the 50 states, plus the District of Columbia, so that $N = 51$, using data releases from September 2015 through September 2024.[1] We analyze these data as annualized quarter-on-quarter growth rates, given that this is both the preferred way that the BEA communicates state GDP data and also because it means we do not contaminate our analysis with benchmark revisions, such as changes in base year or seasonal weights that can lead to level changes (see Aruoba et al. (2016)). The BEA typically publishes four data vintages each year.

We denote the BEA's first estimate of GDP growth in state i in quarter t as y_{it}^{t+1}. The publication lag of this first estimate has changed over time, with the BEA now producing estimates for quarter t toward the end of quarter $(t+1)$; earlier in our sample, this publication lag was longer. Additionally, first estimates have not always been revised in the subsequent quarter. To account for the time variation between the first release and its first revision, we focus on the revision between the first estimate and the estimate produced one year later. Accordingly, we define the four-quarters-ahead revision to the first estimate as:

$$rev_{it}^{4} = y_{it}^{(t+4)+1} - y_{it}^{t+1}. \tag{1}$$

[1] Archived state-level GDP releases are available from https://apps.bea.gov/histdatacore/Regional_Accounts_new.html. We use the seasonally adjusted real GDP estimates from Table SGDP9 in each release.

rev_{it}^4 therefore measures the cumulative revision between the first estimate of state-level GDP growth in quarter t and the most recent estimate available four quarters later, $y_{it}^{(t+4)+1}$.[2] While revisions are an ongoing process, we focus on revisions made during this first year to isolate revisions made primarily due to the arrival of new information rather than benchmark revisions, which the first estimate should not be expected to forecast.

In an online appendix, we also examine the properties of revisions made over the course of 8 and 12 quarters, rev_{it}^8 and rev_{it}^{12}. In general, as summarized below, conclusions about forecast rationality are similar to those reported for the four-quarters-ahead revisions. This said, important revisions are made after one year. In two-thirds of states, the revision bias for rev_{it}^{12} is higher than that for rev_{it}^4.

To compare state-level GDP revisions with revisions to US GDP growth, we align our state-level vintage dataset with monthly data vintages of US GDP obtained from the Real-Time Data Set for Macroeconomists, developed by Croushore and Stark (2001) and maintained by the Federal Reserve Bank of Philadelphia.[3]

3 Rationality of first estimates of state GDP

3.1 (P1): Expected value of revisions

Figure 1 plots a heatmap of the US showing, by state, the bias of the revision, $E(rev_{it}^4)$. Asterisks denote rejection of the null hypothesis of zero bias for the given state, using Newey and West (1987) standard errors with two lags. The figure indicates that bias varies considerably across states, and is statistically significant from zero for 14 states at the 10 percent significance level.[4] Of these 14 states, all but three have negative bias, indicating that their

[2] All but one first release y_{it}^{t+1} in our sample has been revised at least once in the year following its release, with the first estimate for 2022Q2, published in October 2022 and first revised in December 2023, being the only exception. For consistency, we treat that revision as its four-quarters-ahead revision.

[3] See https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/routput.

[4] Of these, the majority also evidence statistically significant bias eight- and twelve-quarters-ahead; see Figures A1 and A2 in the online appendix.

first estimates of state-level GDP were too high. These revision biases are also relatively large, with a statistically significant cross-state bias of -0.24 percentage points.[5] In contrast, the analogous bias to the (aggregate) US GDP growth estimates is about half the size, at 0.11 percentage points, and is not statistically significant (p-value of 0.35).[6]

3.2 (P2): Variance of revisions

Figure 2 shows, by state, the standard deviation of rev_{it}^{4}. The cross-state average standard deviation is 3.15 percent. This compares with a standard deviation of only 0.65 percent for US GDP growth revisions. Every state in Figure 2 sees greater volatility in its revisions than US GDP. The smallest standard deviation, for Ohio at 1.73 percent, is still twice the size of the US estimate. State GDP data are highly uncertain.

While the distribution of state GDP revisions is not perfectly Gaussian (see Figure A3 in the online appendix), assuming normality the standard deviation estimates reported in Figure 2 imply wide confidence intervals around first-release state GDP estimates. With a standard deviation of 3.49, the first estimate of GDP growth in New York, for example, has a 95 percent confidence interval that extends 7 percentage points above and below the first estimate. In real time, this makes it extremely hard to assess, for example, whether the New York economy is contracting.

While state GDP revisions have a higher variance than revisions to US GDP, this is not simply because state GDP growth is more volatile, although it is. When we follow Aruoba (2008) and calculate the "noise-to-signal" ratio for state GDP revisions, we find that the ratio is much higher for all states than it is for national data.[7] Even when measuring relative to

[5]The average bias across the $N = 51$ states (including DC) is: $\overline{rev} = N^{-1}\sum_{i=1}^{N} rev_i = -0.24$, where $rev_i = T^{-1}\sum_{t=1}^{T} rev_{it}^{4}$. Interpreted as a mean group estimator, $\overline{rev}$ is statistically significant from zero at a 1 percent significance level as judged using the variance of the mean group estimator (cf. Pesaran et al. (1996)): $\text{Var}(\overline{rev}) = (N(N-1))^{-1}\sum_{i=1}^{N}(rev_i - \overline{rev})^2$.

[6]While our sample size, T, is too small to permit detailed subsample analysis, to assess if the properties of revisions have changed over time, we note that the revision estimates in this paper are little changed if we exclude GDP estimates for 2020Q2 and 2020Q3, the two quarters most affected by COVID-19.

[7]The noise-to-signal ratio is calculated as the ratio of the standard deviation of four-quarters-ahead

the volatility of the data itself, state GDP estimates are noisier than national estimates.

3.3 (P3): Forecastability of revisions and cross-sectional dependence

To conduct an ex post forecasting exercise to test if revisions to state-level GDP are unpredictable, in the spirit of Aruoba (2008) we estimate three nested models.[8] The least restricted model ("Model 1") regresses a given state's revision, rev_{it}^{4}, on a constant, dummy variables indicating the quarter of the year, Q_t^j, $j = 1, ..., 4$, a linear time trend, the first state-level real GDP growth rate, the latest GDP growth rate for the US, $y_{US,t}^{t+1}$, and lagged year-ahead revisions for the previous quarter for both states and the US. Under efficiency, these coefficients should all equal zero, because all of the conditioning variables are known at the time the first estimate of state-level GDP, y_{it}^{t+1}, is published. We include US data in the efficiency regression to test for cross-state dependencies driven by a US "common factor."

To isolate the in-sample predictive content of the US data, we test "Model 1" against "Model 2," which drops the US data. Finally, to test efficiency relative to all variables in Model 1, we test whether all the coefficients in Model 1 (including the constant term) equal zero.[9] When we reject the null hypothesis that all coefficients are zero, we reject the "news" hypothesis.

From Figure 3, we infer that for 25 states one could "improve" (at a 1 or 5 percent significance level) the first estimate of state GDP by conditioning on the (known) first estimate of US GDP. For these states, their revisions are cross-sectionally correlated and driven by movements in the US data. In turn, Figure 4 evidences that, using the expanded set of variables,

revisions to a given state's GDP, SD $\left(rev_{it}^{4}\right)$, and the standard deviation of the revised growth rates, SD $\left(y_{it}^{(t+4)+1}\right)$. See Figure A6 in the online appendix for estimates, by state, of the noise-to-signal ratio.

[8] Future research, as the sample of state-level GDP data increases, should repeat this as a real-time (ex ante) forecasting exercise.

[9] We use autocorrelation robust F-tests, acknowledging that T is quite small. Robustness checks confirm that our main conclusions hold when we do not use Newey and West (1987) standard errors and use an LR test; see Figures A9 and A10.

again all known at the time the first estimate of state GDP was made, all but four states have predictable revisions (tested at a 1 or 5 percent significance level). These four states are Arkansas, Colorado, Georgia, and Washington. All states, save Georgia, evidence inefficiency at the 10 percent level.[10] Looking back to Figures 1 and 3, and drawing on the robustness exercises in the online appendix, we see that only Georgia appears to have "well behaved" revisions (at the 10 percent significance level) across the battery of tests, (P1) through (P3).

To also exploit the cross-sectional dimension when testing (P3), finally we estimate a panel data variant of these efficiency regressions. This regression pools the coefficients on variables across states, allowing for state fixed effects: $rev_{it}^4 = \alpha_i + \Sigma_{j=1}^4 \lambda_j Q_t^j + \tau t + \beta y_{it}^{t+1} + \gamma y_{US,t}^{t+1} + \delta rev_{i,t-1}^4 + \psi rev_{US,t-1}^4 + \epsilon_{it}$. The null hypothesis of forecast efficiency, pooled across N and T, $H_0 : \lambda_j = \tau = \beta = \gamma = \delta = \psi = 0$, is again rejected, with a p-value of 0.000.

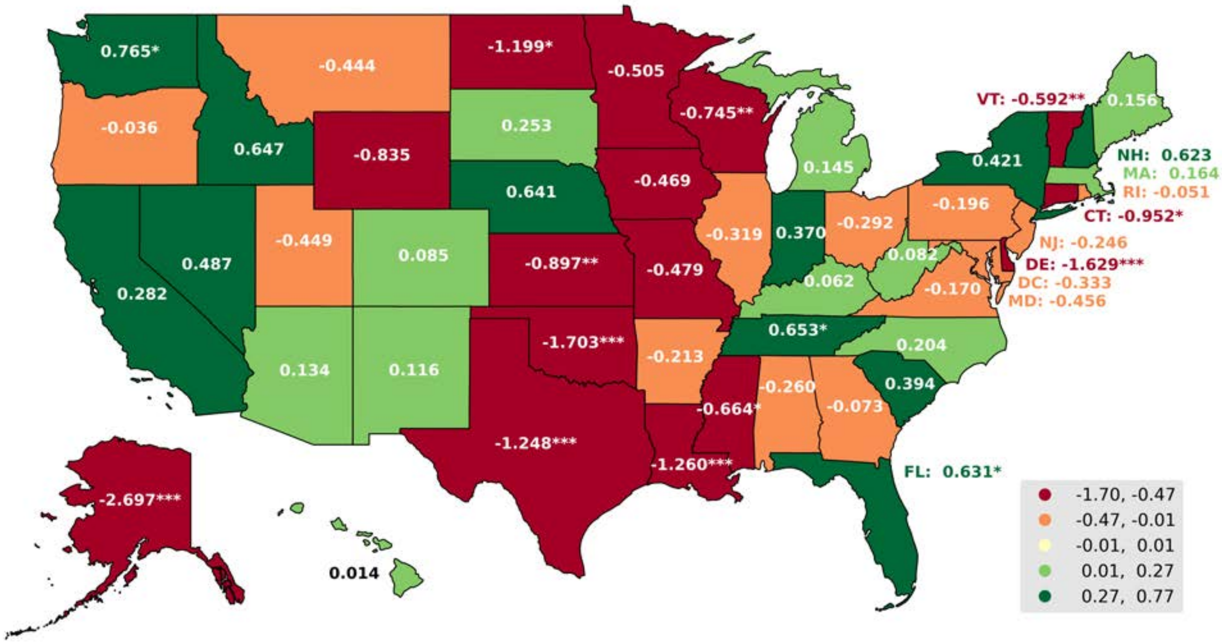
4 Conclusion

This paper finds that revisions to state-level GDP data are not well behaved, except for Georgia. Users of state-level GDP growth data should be aware that first releases tend to be biased, have a large variance, and that revisions can be predicted including by conditioning on the latest US estimates.

This paper also evidences considerable heterogeneity across states in terms of how well behaved their GDP data are. Future research might explore the drivers of this heterogeneity. In an illustrative exercise in this direction, we find that states with better behaved revisions, at least according to (P2), have regional economies that more closely resemble the structure of the US as a whole. Specifically, we find a statistically significant negative Spearman rank correlation between $SD(rev_{it}^4)$ and the Hachman measure of state employment diversity relative to the US, as used in Bokun et al. (2023).

[10] Colorado and Washington remain efficient, at the 10 percent level, when autocorrelation-robust standard errors are not used; cf. Figure A10 in the appendix.

Figure 1: (P1): Mean revision by state ($E(rev_{it}^4)$), 2015Q1-2023Q2



Notes: Quarterly growth at annualized rates. Asterisks denote rejection of the null hypothesis of zero bias for the given state, using Newey and West (1987) standard errors with two lags, at the 1 (***), 5 (**), and 10 (*) percent significance levels.

Figure 2: (P2): Standard deviation of revisions by state, ($SD(rev_{it}^4)$), 2015Q1-2023Q2

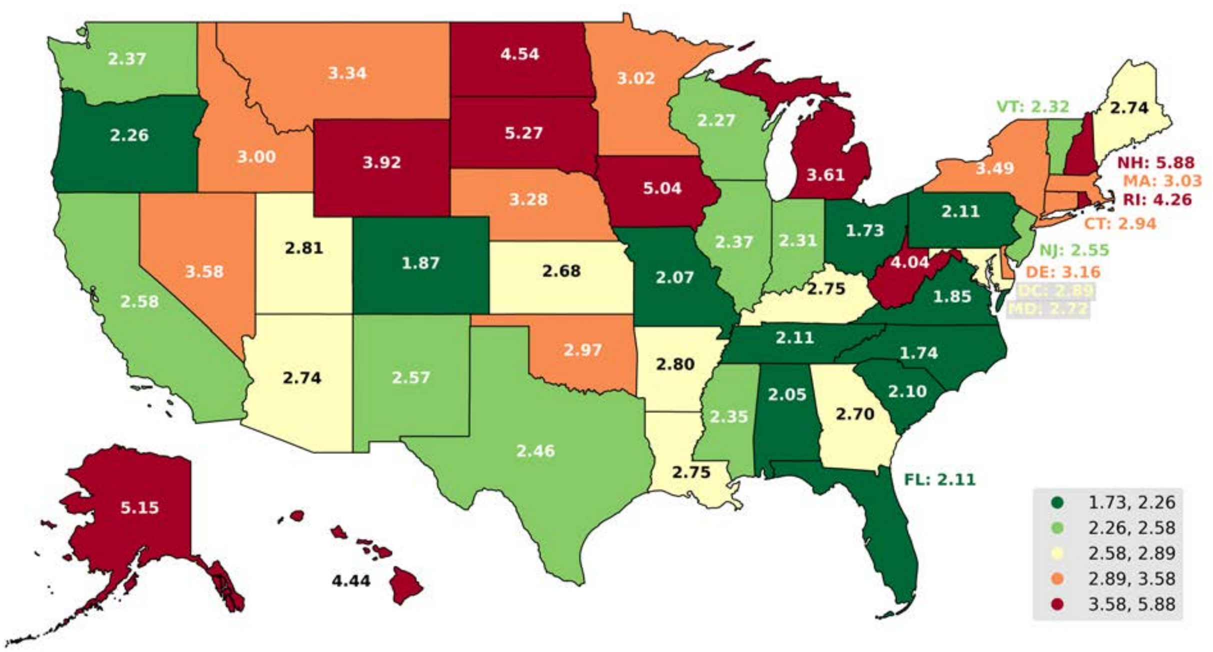
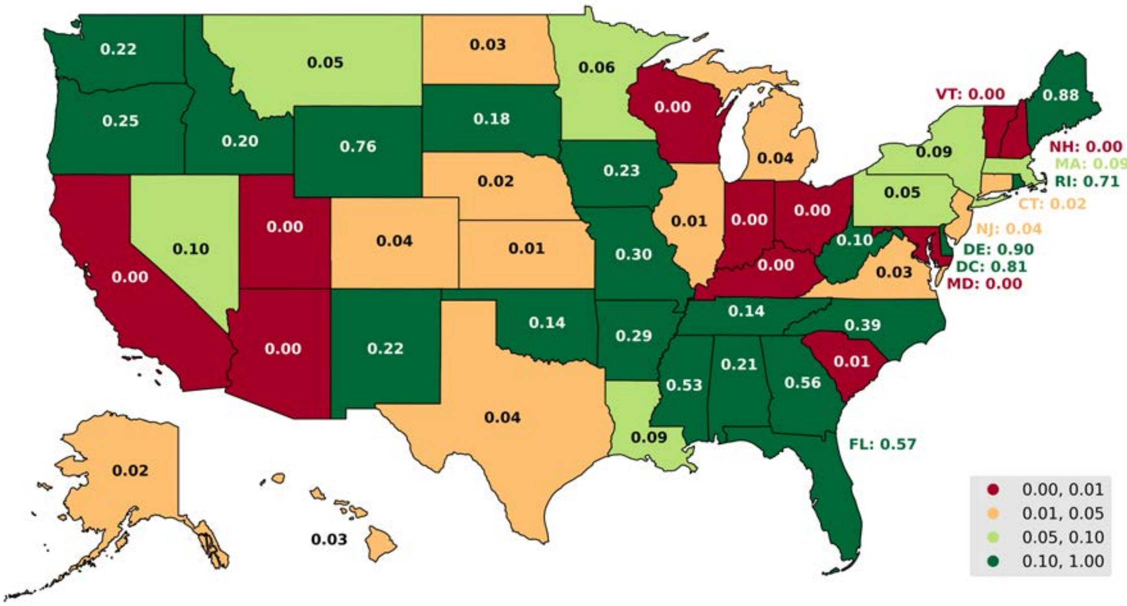
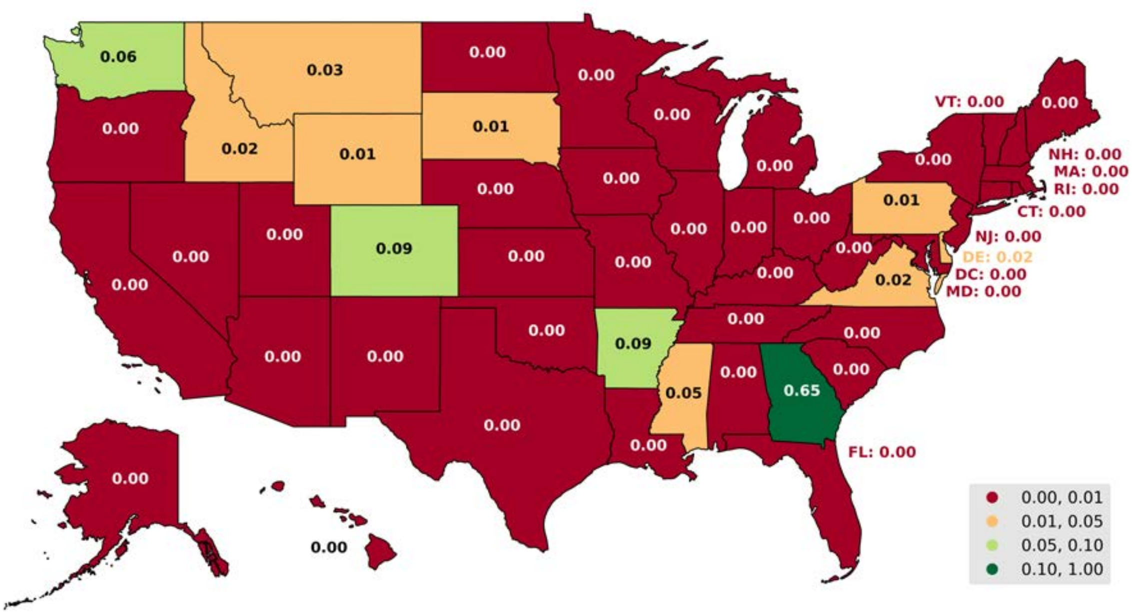


Figure 3: (P3): Efficiency of data revisions by state: Do known US data explain state GDP revisions?



Notes: P-values, using Newey and West (1987) standard errors with two lags, testing $H_0 : \gamma_i = \psi_i = 0$ in the regression: $rev^4_{it} = \alpha_i + \Sigma^4_{j=1}\lambda_{ij}Q^j_t + \tau_i t + \beta_i y^{t+1}_{it} + \gamma_i y^{t+1}_{US,t} + \delta_i rev^4_{i,t-1} + \psi_i rev^4_{US,t-1} + \epsilon_{it}$.

Figure 4: (P3): Testing the efficiency of data revisions by state



Notes: P-values, using Newey and West (1987) standard errors with two lags, testing whether all the coefficients in the following regression equal zero: $rev^4_{it} = \alpha_i + \Sigma^4_{j=1}\lambda_{ij}Q^j_t + \tau_i t + \beta_i y^{t+1}_{it} + \gamma_i y^{t+1}_{US,t} + \delta_i rev^4_{i,t-1} + \psi_i rev^4_{US,t-1} + \epsilon_{it}$.

References

Arias, Maria A., Charles S. Gascon, and David E. Rapach (2016). “Metro business cycles.” *Journal of Urban Economics*, 94, pp. 90–108. doi:10.1016/j.jue.2016.05.005.

Aruoba, S. Borağan (2008). “Data revisions are not well behaved.” *Journal of Money, Credit and Banking*, 40, pp. 319–340. doi:10.1111/j.1538-4616.2008.00115.x.

Aruoba, S. Borağan, Francis X. Diebold, Jeremy Nalewaik, Frank Schorfheide, and Dongho Song (2016). “Improving GDP measurement: a measurement-error perspective.” *Journal of Econometrics*, 191, pp. 384–397. doi:10.1016/j.jeconom.2015.12.009.

Baumeister, Christiane, Danilo Leiva-León, and Eric Sims (2024). “Tracking Weekly State-Level Economic Conditions.” *Review of Economics and Statistics*, 106(2), pp. 483–504. doi:10.1162/rest_a_01171.

Bokun, Kathryn O., Laura E. Jackson, Kevin L. Kliesen, and Michael T. Owyang (2023). “FRED-SD: A real-time database for state-level data with forecasting applications.” *International Journal of Forecasting*, 39(1), pp. 279–297. doi:10.1016/j.ijforecast.2021.11.008.

Croushore, Dean (2011). “Frontiers of real-time data analysis.” *Journal of Economic Literature*, 49, pp. 72–100. doi:10.1257/jel.49.1.72.

Croushore, Dean and Tom Stark (2001). “A real-time data set for macroeconomists.” *Journal of Econometrics*, 105(1), pp. 111–130. doi:10.1016/S0304-4076(01)00072-0.

Galvão, Ana Beatriz and James Mitchell (2023). “Real-time perceptions of historical GDP data uncertainty.” *Oxford Bulletin of Economics and Statistics*, 85(3), pp. 457–481. doi:10.1111/obes.12542.

Hamilton, James D. and Michael T. Owyang (2012). “The propagation of regional recessions.” *Review of Economics and Statistics*, 94(4), pp. 935–947. doi:10.1162/REST_a_00197.

Mankiw, N. Gregory and Matthew D. Shapiro (1986). “News or noise: An analysis of GNP revisions.” *Survey of Current Business*, pp. 20–25. doi:10.3386/w1939.

Nakamura, Emi and Jón Steinsson (2014). “Fiscal stimulus in a monetary union: Evidence from US regions.” *American Economic Review*, 104(3), p. 753–92. doi:10.1257/aer.104.3.753.

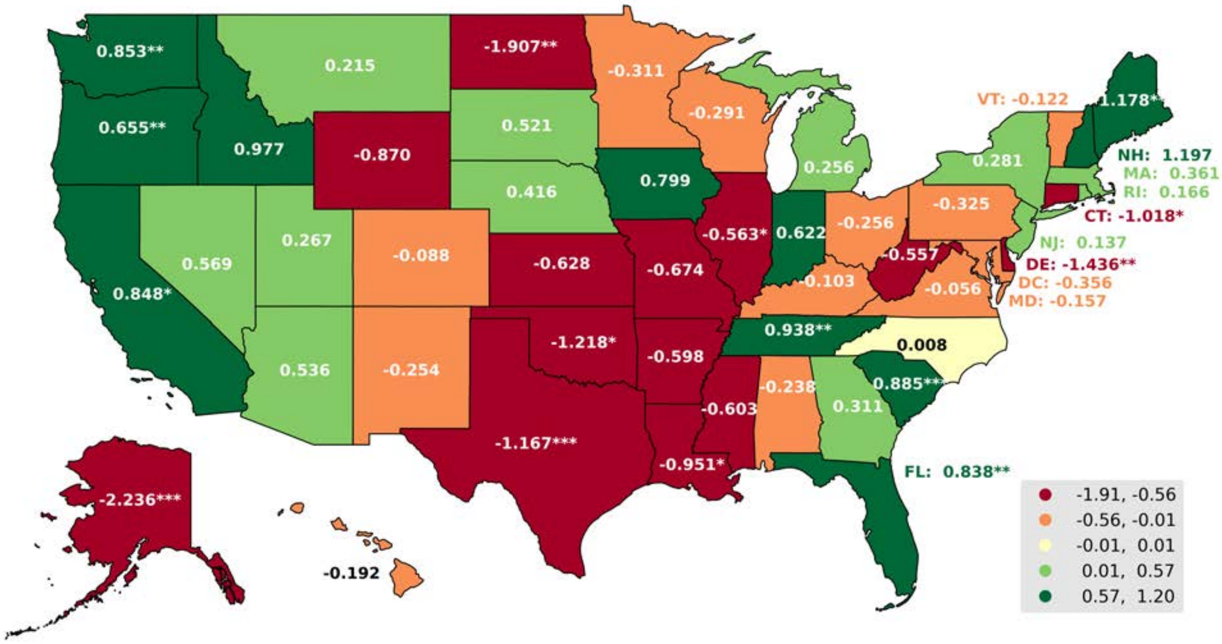
Newey, Whitney K. and Kenneth D. West (1987). “A simple positive semi-definite heteroskedasticity and autocorrelation-consistent covariance matrix.” *Econometrica*, 55, pp. 703–708. doi:10.2307/1913610.

Pesaran, M. Hashem, Ron Smith, and Kyung So Im (1996). “Dynamic linear models for heterogenous panels.” In László Mátyás and Patrick Sevestre, editors, *The Econometrics of Panel Data: A Handbook of the Theory with Applications*, pp. 145–195. Springer Netherlands, Dordrecht. doi:10.1007/978-94-009-0137-7_8.

A Online Appendix for "Are Revisions to State-Level GDP Data in the US Well Behaved?" by James Mitchell and Taylor Shiroff

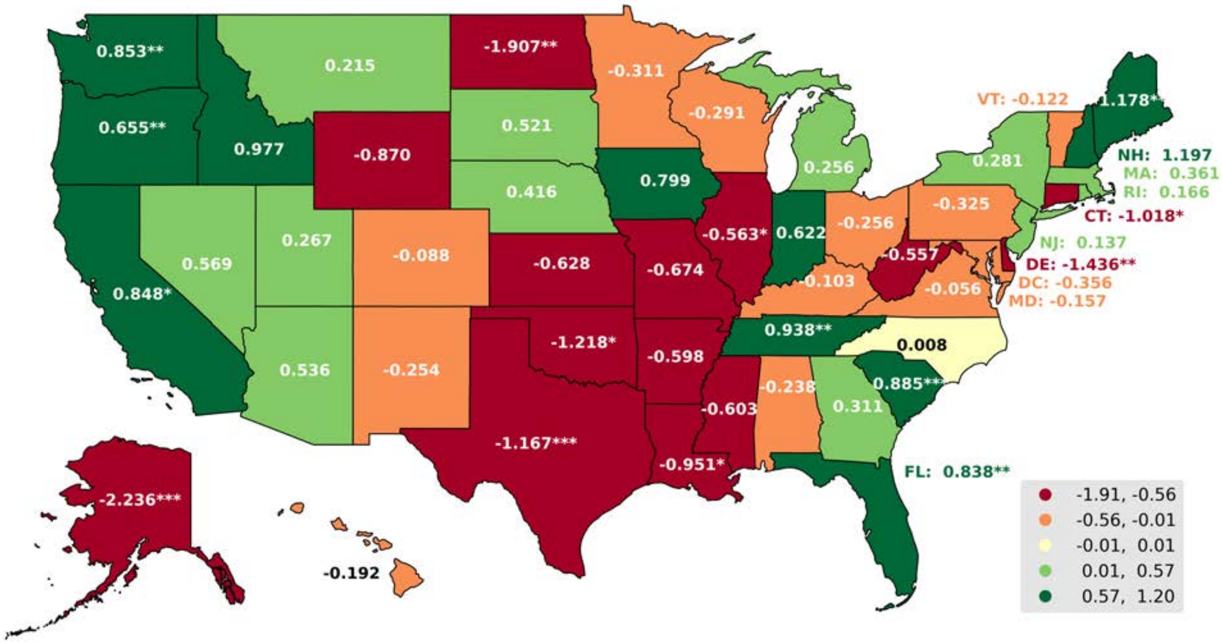
This appendix contains supplementary empirical results referred to in the main paper.

Figure A1: (P1): Mean revision by state eight-quarters-ahead, $E(rev_{it}^{8})$



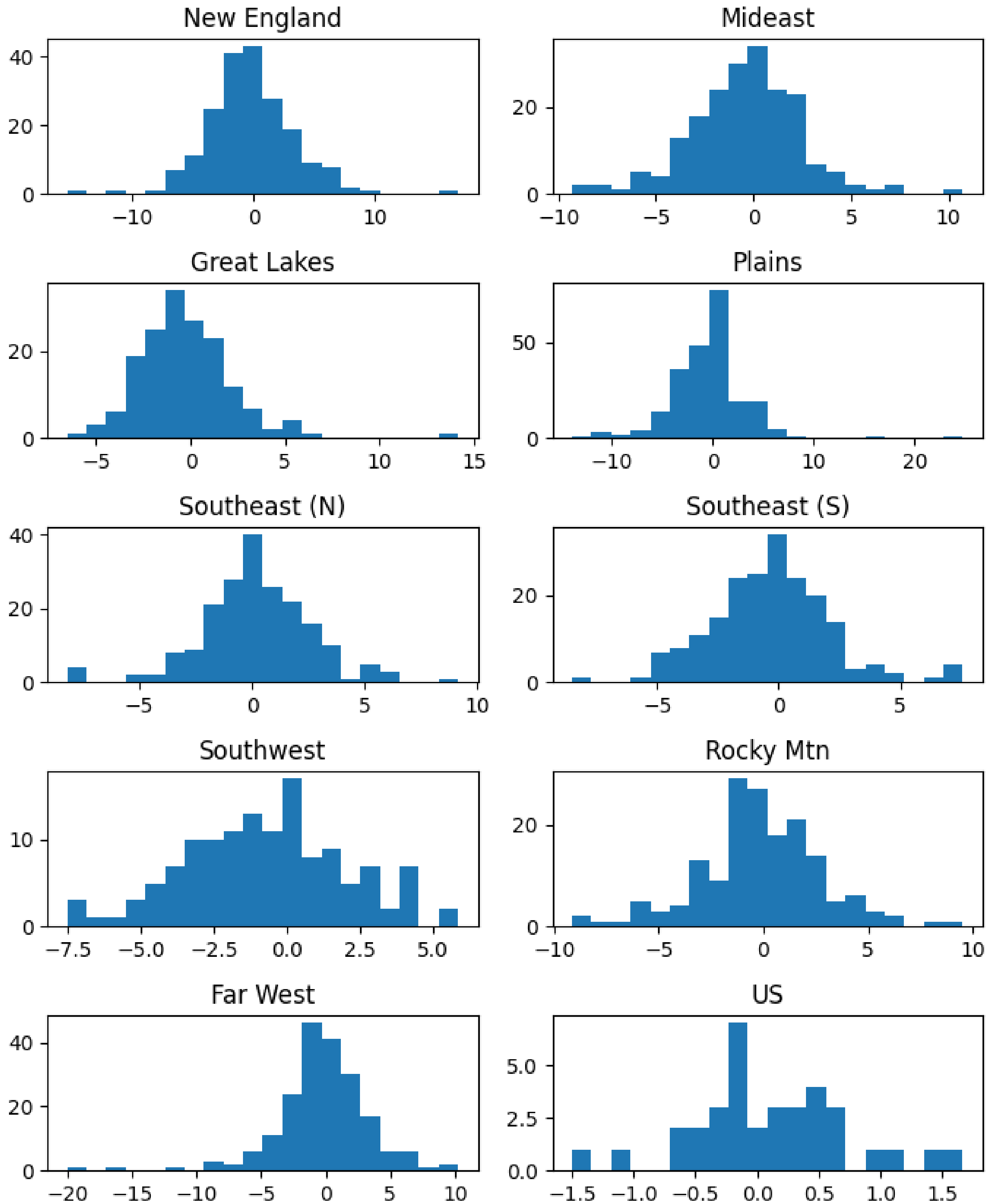
Notes: Quarterly growth at annualized rates. Sample: first estimates of GDP growth for 2015Q1 through 2022Q2. Asterisks denote rejection of the null hypothesis of zero bias for the given state, using Newey and West (1987) standard errors with three lags, at the 1 (***), 5 (**), and 10 (*) percent significance levels.

Figure A2: (P1): Mean revision by state twelve-quarters-ahead, $E(rev_{it}^{8})$



Notes: Quarterly growth at annualized rates. Sample: first estimates of GDP growth for 2015Q1 through 2021Q2. Asterisks denote rejection of the null hypothesis of zero bias for the given state, using Newey and West (1987) standard errors with three lags, at the 1 (***), 5 (**), and 10 (*) percent significance levels.

Figure A3: Histograms of four-quarters-ahead revisions by BEA region (equal weighting each state within a region) and for the US as a whole



Notes: New England: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. Mideast: Delaware, Maryland, New Jersey, New York, Pennsylvania, and Washington, D.C.. Great Lakes: Illinois, Indiana, Michigan, Ohio, and Wisconsin. Plains: Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota. We break Southeast into north (N): North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. Southeast south (S) is then assumed to comprise: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, and Mississippi. Southwest: Arizona, New Mexico, Oklahoma, and Texas. Rocky Mountain: Colorado, Idaho, Montana, Utah, and Wyoming. Far West: Alaska, California, Hawaii, Nevada, Oregon, and Washington.

Figure A4: (P2): Standard deviation of eight-quarters-ahead revisions by state, $SD(rev_{it}^{8})$

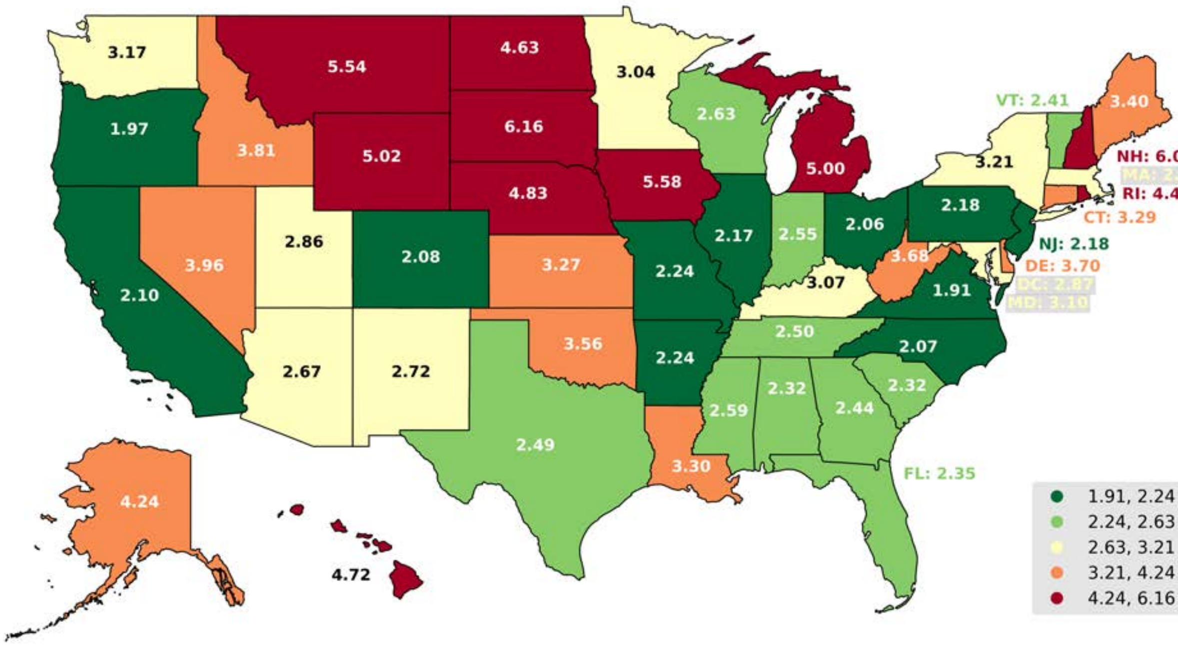


Figure A5: (P2): Standard deviation of twelve-quarters-ahead revisions by state, $SD(rev_{it}^{12})$

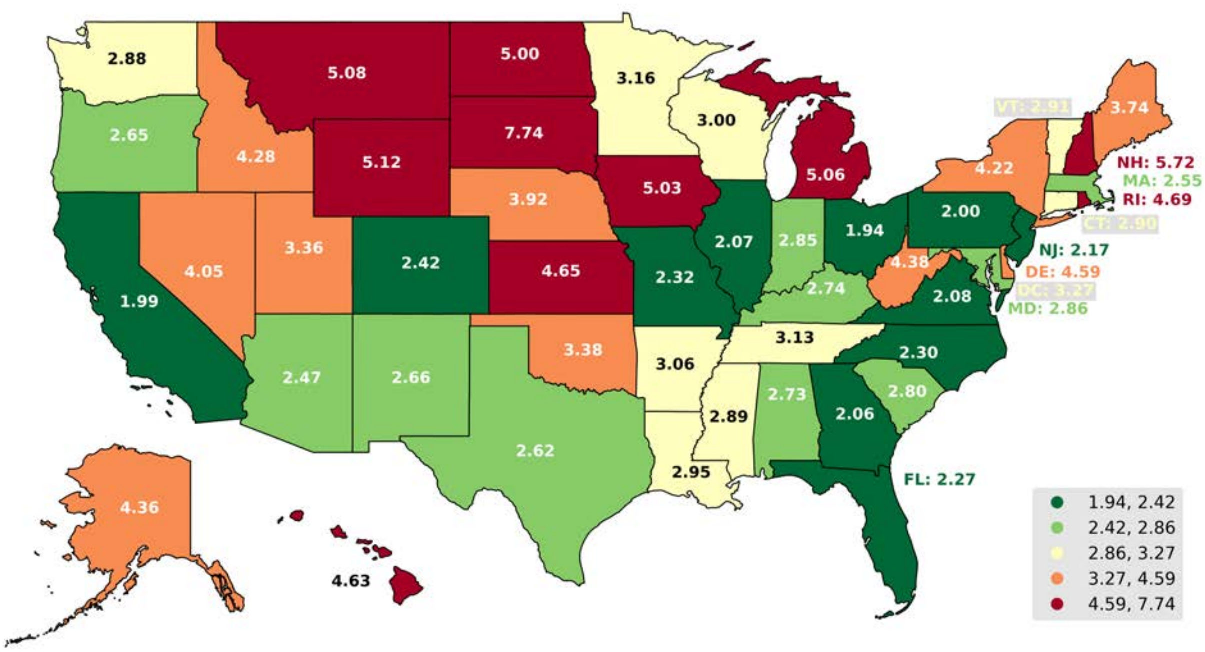
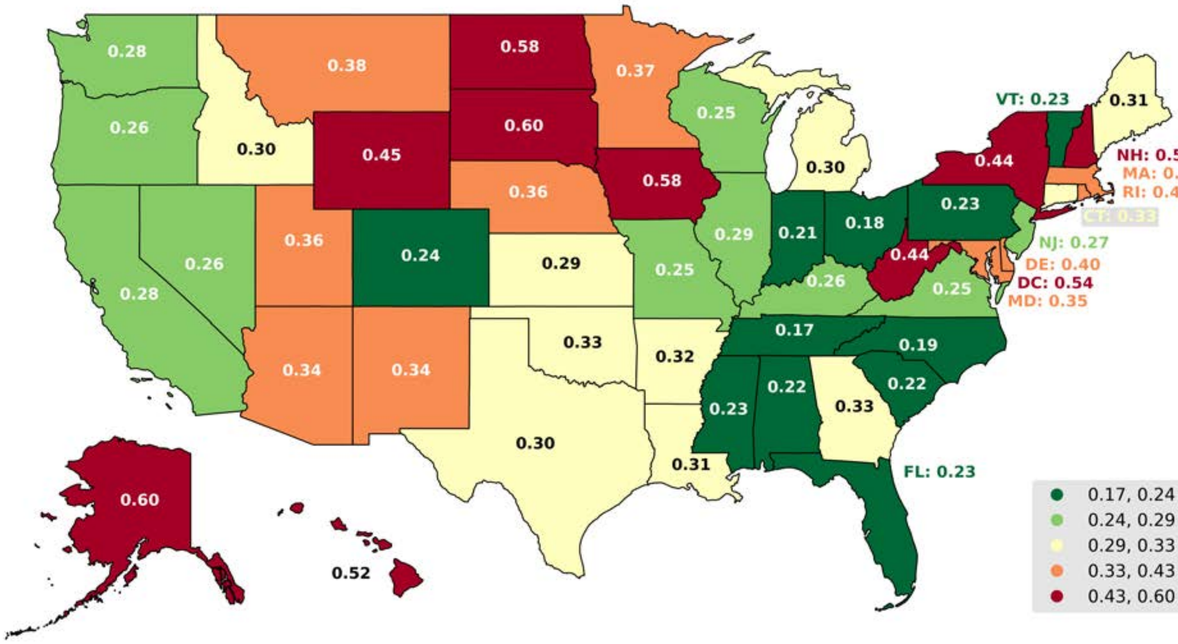


Figure A6: Noise-to-signal ratios by state for rev_{it}^{4}



Notes: The noise-to-signal ratio is calculated as the ratio of the standard deviation of four-quarters-ahead revisions to a given state's GDP, SD $\left(rev_{it}^{4}\right)$, and the standard deviation of the revised growth rates, SD $\left(y_{it}^{(t+4)+1}\right)$.

Figure A7: Noise-to-signal ratios by state for rev_{it}^{8}

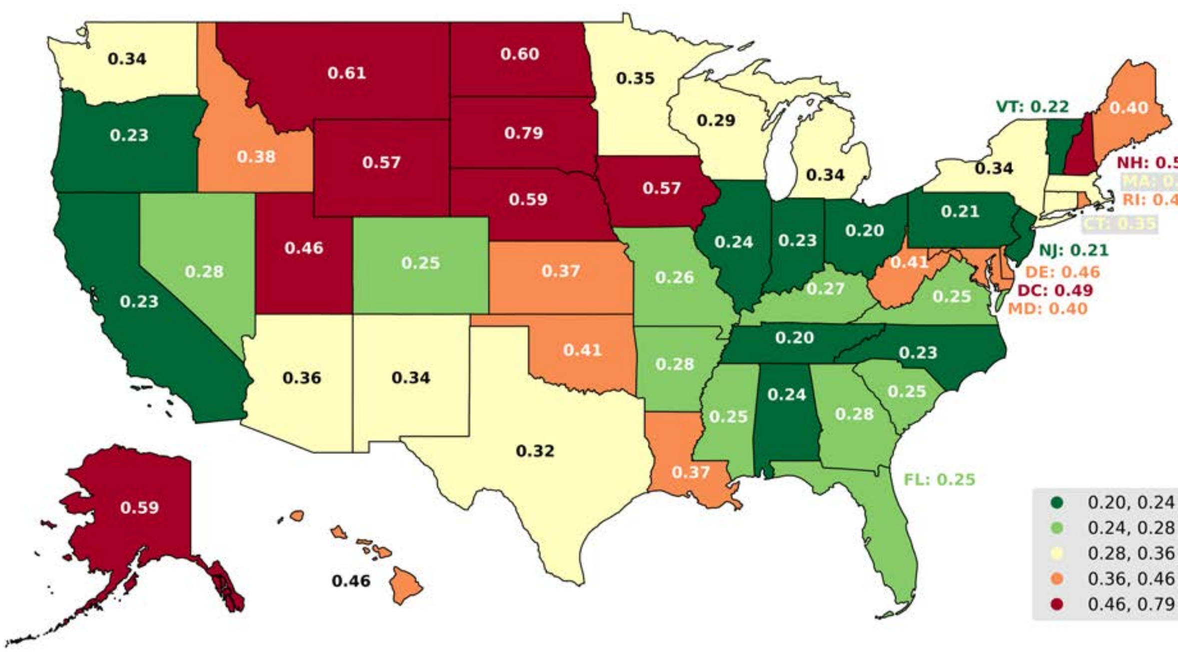


Figure A8: Noise-to-signal ratios by state for rev_{it}^{12}

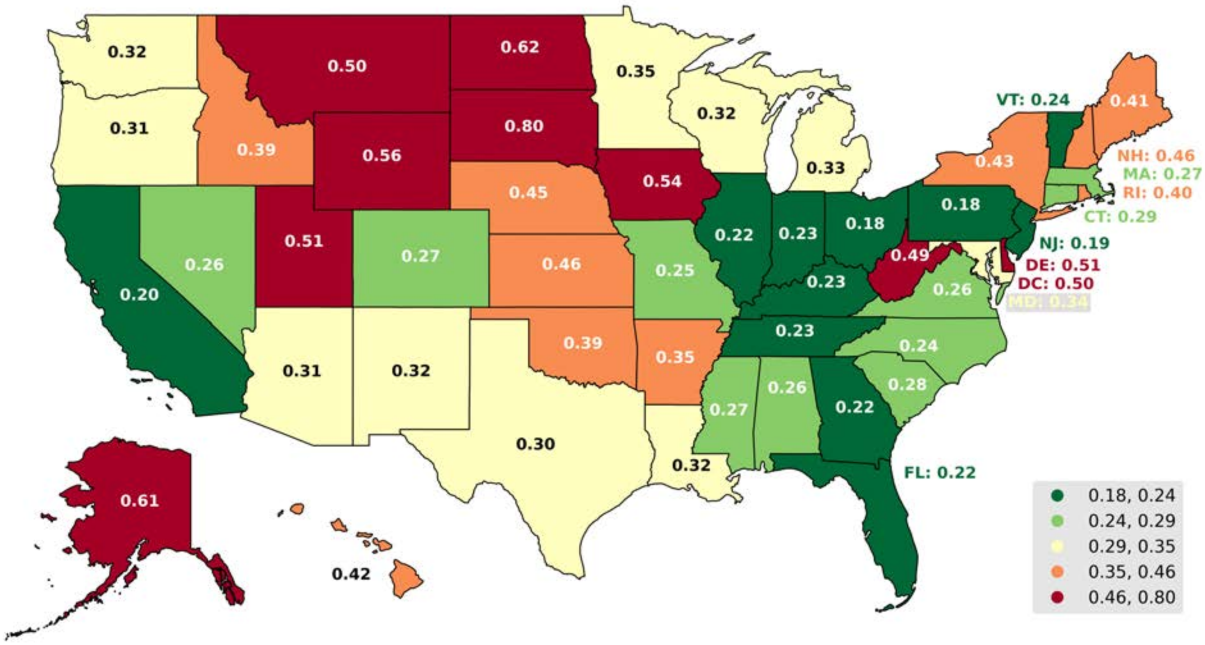
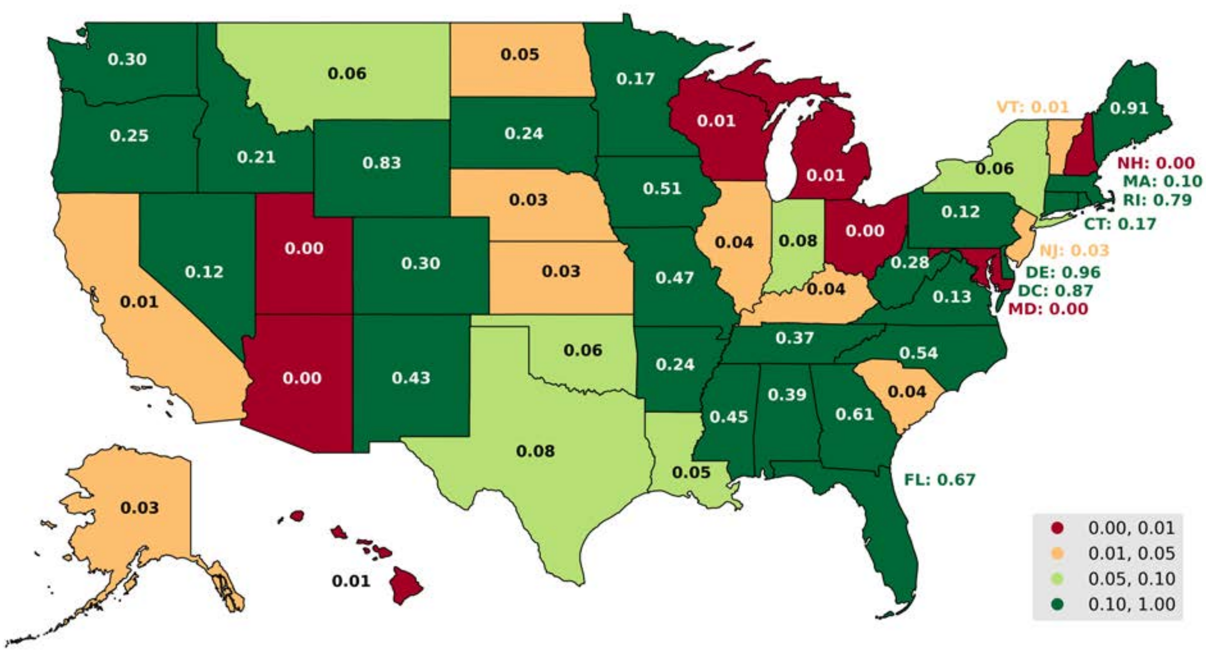
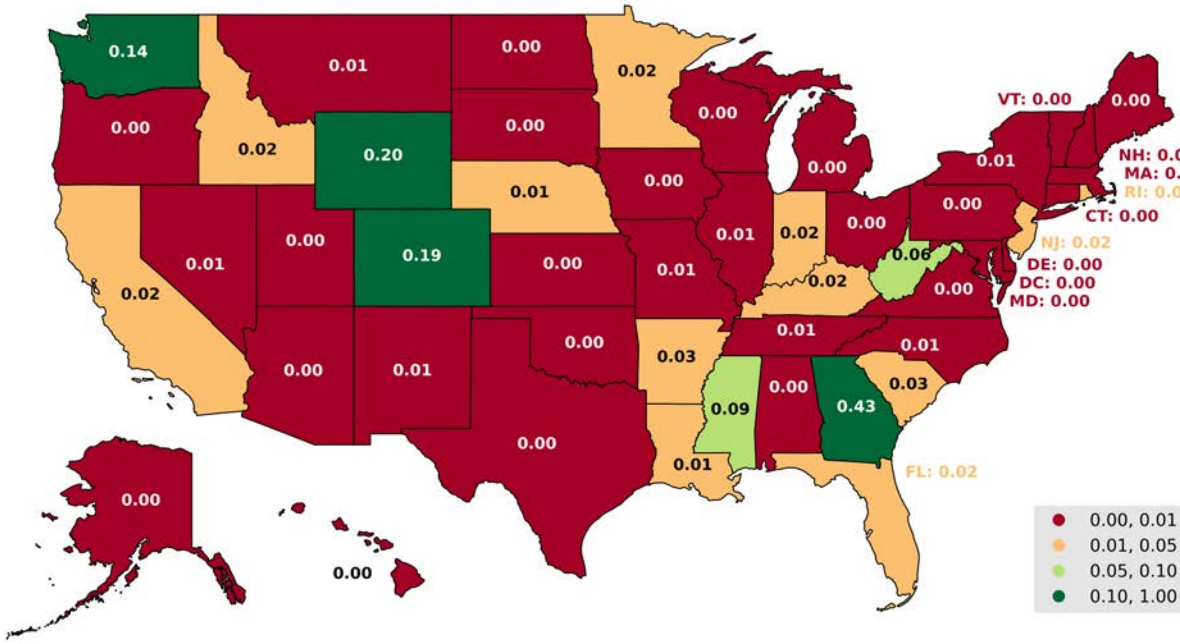


Figure A9: (P3): LR tests of the efficiency of data revisions by state: Do known US data explain revisions to state GDP?



Notes: P-values testing $H_0 : \gamma_i = \psi_i = 0$ in the regression: $rev_{it}^4 = \alpha_i + \Sigma_{j=1}^4 \lambda_{ij} Q_t^j + \tau_i t + \beta_i y_{it}^{t+1} + \gamma_i y_{US,t}^{t+1} + \delta_i rev_{i,t-1}^4 + \psi_i rev_{US,t-1}^4 + \epsilon_{it}$.

Figure A10: (P3): LR tests of the efficiency of data revisions by state



Notes: P-values testing whether all the coefficients in the following regression equal zero: $rev_{it}^{4} = \alpha_i + \Sigma_{j=1}^{4} \lambda_{ij} Q_t^j + \tau_i t + \beta_i y_{it}^{t+1} + \gamma_i y_{US,t}^{t+1} + \delta_i rev_{i,t-1}^{4} + \psi_i rev_{US,t-1}^{4} + \epsilon_{it}$.