REVISITING THE PREDICTION OF GDP GROWTH USING AR MODELS

DANIEL LEWIS

Date: May 2, 2023.

1. Project motivation

In this essay we revisit the topic of predicting GDP growth using autoregressive (AR) models, which the author considered in [Lew22].

There, we analyzed a dataset consisting of Gross Domestic Product (GDP) figures for each of the 50 US states over the last 17 years, and discussed whether the AR(4) model provided the best predictions, as theory suggests it should. A further investigation of *loc. cit.* contemplated the counterfactual in which the 2020 Covid-19 pandemic had not occurred, quantifying how great the disparity was between forecasted growth and actual growth post-slowdown.

We now reassess our methodology, equipped with 2022 Q3 GDP data that was not available when the project was first proposed. We ask once again if AR(4) predictions are best, or if, as then, the AR(16) model performs better in practice. Moreover, we compare these AR specifications against, first, a constant growth model, then, second, a weighted average of AR(4) and constant growth, which may perform better still. Finally, as an application, we return to the hypothetical scenario in which the 2020 Covid-19 pandemic had not occurred.

2. Outline

In Section 3 we describe our dataset.

Section 4 outlines the models we are to compare.

Section 5 presents the performance metrics we use in our analysis: naïve forecasting error and mean squared forecasting error.

We present our explanatory data analysis in Section 6 and our plan of analysis in Section 7. This includes consideration of nonparametric identification, weak instruments, natural choice of bounds, and whether bounds can be used as a robustness check.

Finally, we describe our results in Section 8 and our conclusions in Section 9.

We once again include three appendices: the first, Appendix I, describes the reproducibility of our results; the second, Appendix II, presents the graphical results of our exploratory data analysis. while Appendix III considers our "side quest" of imagining economic growth had the 2020 Covid-19 pandemic not occurred.

2.1. **Acknowledgments.** Many thanks to Tiemen Woutersen for suggesting this line of inquiry.

3. The dataset

The data used in this project was once again obtained from the Federal Reserve Bank of St. Louis [FRE22]. The dataset consists of quarterly data on the seasonally adjusted annual rate of the all industry GDP total, measured in millions of dollars, for the state of Arizona. The data stretches from 2005 Q1 to 2022 Q3: 71 quarters in all (one more than in [Lew22]).

4. Our models

Recall that the linear model works well for continuous outcomes if we have only a small number of regressors, such as when predicting GDP growth.

The simplest autoregressive (AR) model considers the regression of Y_t on $\{1, Y_{t-1}\}$:

$$Y_t = \alpha + \rho_1 Y_{t-1} + \epsilon_t. \tag{AR(1)}$$

Theory suggests we may prefer to regress Y_t on $\{1, Y_{t-1}, Y_{t-2}, Y_{t-3}\}$:

$$Y_t = \alpha + \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + \rho_3 Y_{t-3} + \rho_4 Y_{t-4} + \epsilon_t, \tag{AR(4)}$$

while we saw in [Lew22] the unexpected superiority of the model

$$Y_t = \alpha + \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + \cdots + \rho_{16} Y_{t-16} + \epsilon_t, \tag{AR(16)}$$

In each case, we produce Ordinary Least Squares (OLS) predictions of the coefficients α and ρ_i iteratively.

The constant growth model gives an alternative prediction $\tilde{Y}_{T+1}$ via

$$\tilde{Y}_{T+1} = \hat{\kappa} Y_T, \tag{constant}$$

where we estimate the growth rate, κ, as the mean $\hat{\kappa} = \frac{1}{T} \sum_{t=1}^{T-1} Y_{t+1}/Y_t$.

We also consider a third model taking a simple average of the AR(4) and constant growth predictions:

$$\check{Y}_{T+1} = \frac{\hat{Y}_{T+1} + \tilde{Y}_{T+1}}{2}. \tag{hybrid}$$

This project aims to investigate whether the AR(4) model outperforms these other specifications when applied to our real world dataset (described in Section 3), as empirical evidence suggests it should.

5. Performance metrics

We use the same metrics to evaluate prediction performance as in [Lew22]: *Mean Squared Forecasting Error (MSFE)*

$$\text{MSFE}(Y_t, \tilde{Y}_t) := E[(Y_t - \tilde{Y}_t)^2],$$

and a naïve "forecasting error" (FE),

$$\text{FE} = Y_t - \tilde{Y}_t.$$

We seek the lowest MSFE and the lowest FE in absolute value.

6. Exploratory data analysis

In [Lew22], we computed AR(4) predictions of Arizona's GDP for 2022 Q3, Q4 and 2023 based on the preexisting data. Our first task is to compare the newly available true 2022 Q3 figure against our prior predictions. The plot is included in Appendix II.

We also plot a comparison of the true growth figures against the straight line provided by a constant growth model. Note (see Figure 5) that the constant growth model underestimates growth prior to the 2008/09 Great Recession and typically overestimates subsequent growth. The discrepancy was greatest at the height of the 2020 Covid-19 pandemic. However, GDP growth in the state of Arizona now appears to be rising faster than the constant growth model predicts.

7. Plan of analysis

The first step in our analysis is to train each model on the first 70 datapoints, and compare (via naïve FE) the predicted GDP figures for 2022 Q3 against the true figures now included in our dataset.

Our second step is to train each model on the first 50 datapoints in our sample, and use this to make predictions for the next 21 datapoints. Thus we may compute an estimate of the MSFE for these predictions for each model.

Finally, we address our "side quest", comparing predicted growth in 2020 against the actual figures marking the downturn.

7.1. **Nonparametric identification.** The identification strategy for AR(p) models is classical, due to Yule [Yul27] and Walker [Wal31].

The Yule–Walker equations take the form

$$\gamma_m = \sum_{k=1}^{p} \rho_k \gamma_{m-k} + \sigma_\epsilon^2 \delta_{m,0},$$

where $m = 0, \ldots, p$. Here γ_m is the *autocovariance function* $\gamma_m = \mathrm{Cov}(Y_{t_m}, Y_{t_{m+1}})$ of Y_t, σ_ϵ is the standard deviation of the noise term ϵ_t in equation (AR(4)), and $\delta_{m,0}$ is the Kronecker delta function. That is, $\delta_{m,0} = 1$ when $m = 0$ and equals zero otherwise. In matrix form,

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \vdots \\ \gamma_p \end{pmatrix} = \begin{pmatrix} \gamma_0 & \gamma_{-1} & \gamma_{-2} & \cdots \\ \gamma_1 & \gamma_0 & \gamma_{-1} & \cdots \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \cdots \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_p \end{pmatrix}, \tag{7.1}$$

and, additionally,

$$\gamma_0 = \sum_{k=1}^{p} \rho_k \gamma_{-k} + \sigma_\epsilon^2. \tag{7.2}$$

Yule and Walker's idea is to first solve the system (7.1) for $\rho_1, \ldots, \rho_p$, then solve (7.2) for σ_ϵ^2, thus identifying all coefficients in the system.

Example 7.1 (Identification of an AR(4) model)**.** In the special case $p = 4$ of interest to us, the setup described above simplifies to the system of five equations:

$$\begin{cases} \gamma_0 = \sum_{k=1}^4 \rho_k \gamma_{-k} + \sigma_\epsilon^2, \\ \gamma_1 = \sum_{k=1}^4 \rho_k \gamma_{1-k}, \\ \gamma_2 = \sum_{k=1}^4 \rho_k \gamma_{2-k}, \\ \gamma_3 = \sum_{k=1}^4 \rho_k \gamma_{3-k}, \\ \gamma_4 = \sum_{k=1}^4 \rho_k \gamma_{4-k}. \end{cases}$$

Thus, we are to solve the matrix equation

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix} = \begin{pmatrix} \gamma_0 & \gamma_{-1} & \gamma_{-2} & \gamma_{-3} \\ \gamma_1 & \gamma_0 & \gamma_{-1} & \gamma_{-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \gamma_{-1} \\ \gamma_3 & \gamma_2 & \gamma_1 & \gamma_0 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{pmatrix},$$

together with the equation for γ_0. Since autocovariance functions satisfy $\gamma_k = \gamma_{-k}$, the matrix of autocovariance functions is symmetric and so may be inverted to solve the system:

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{pmatrix} = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 \\ \gamma_1 & \gamma_0 & \gamma_1 & \gamma_2 \\ \gamma_2 & \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_3 & \gamma_2 & \gamma_1 & \gamma_0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix}$$

Finally, one reads off

$$\sigma_\epsilon^2 = \gamma_0 - \sum_{k=1}^{p} \rho_k \gamma_{-k}.$$

7.2. **Weak instruments.** We do not apply an instrumental variables framework, so there is no need to discuss the possibility of weak instruments in our setup.

7.3. **Natural choice of bounds.** Since the trend in GDP is increasing, if this trend continues then the estimate of GDP in any particular quarter should be at least that of the previous quarter. Provided the estimated growth rate $\hat{\kappa} \geq 0$, then the constant growth model satisfies this requirement. For our other model specifications, whether or not this criterion is satisfied is best assessed visually.

7.4. **Bounds as robustness check.** Again, seeking bounds to verify the robustness of any particular model is best assessed visually. One should note the inaccuracy of long-run projections, but be confident in the short-term.

One potential avenue for checking robustness of the models in our setup is to take the constant growth predictions as a baseline. One may expect AR(p) predictions to exceed this baseline, hopefully taking into account further GDP growth due to innovation above and beyond that predicted by the constant growth parameter estimated based upon economic performance in previous quarters. Of course, this may fail in the event of an economic downturn.

Model	Forecast ($m)	FE ($m)	Runtime (s)
True GDP	463653.7	-	-
AR(4)	461270.2	-2383.544	0.76
Const.	467882.1	4228.378	0.83
Hybrid	464576.1	922.417	2.21

Table 8.1. Step 1: we train each model on the first 70 datapoints in our sample and predict growth in 2022 Q3.

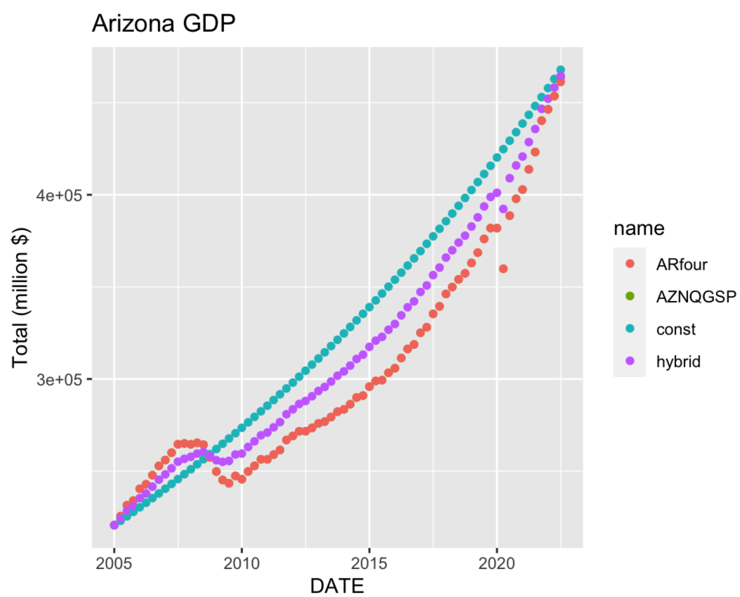

Figure 1. Step 1: note that the first 70 'AZNQGSP' (true GDP) datapoints overlap with the ARfour predictions trained upon them.

8. Results

Recall that our first-stage analysis is to train each model on the first 70 datapoints, and compare the predicted GDP figures for 2022 Q3 against the true figures. The results are provided in Table 8.1 and Figure 1.

Note in particular how the AR(4) model, while underestimating true GDP growth, outperforms the constant growth model, which overestimates growth by a larger amount. The hybrid model, as an average of the two, performs the best, overestimating growth by a reduced amount (albeit with greater runtime).

Step two of the analysis consists of training each model on the first 50 datapoints in the sample, using this to predict the next 21 datapoints. The results are provided in Table 8.2 and Figures 2 and 3.

This time one should note how the AR(4) model is outperformed by the constant growth specification, which performs best, while the hybrid model fits the data better than AR(4).

In light of the unexpected result of [Lew22] in which the AR(16) specification outperformed that of AR(4), we find it informative to also consider the AR(16)

Model	MSFE	Runtime (s)
AR(4)	1522965460	0.69
Const.	880893619	0.94
Hybrid	1176194559	1.90
AR(16)	493189423	1.07
Hybrid16	670115229	2.32

TABLE 8.2. Step 2: we train each model on the first 50 datapoints in our sample and predict the remaining 21 datapoints.

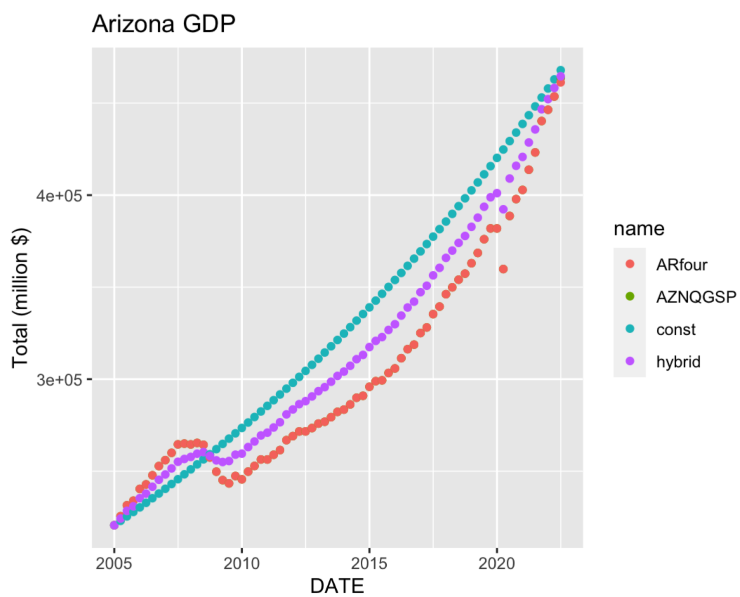

FIGURE 2. Step 2: note how the constant growth model outperforms AR(4) in this case.

model and its hybrid (a simple average of AR(16) and constant growth predictions), denoted Hybrid16 in Table 8.2.

It is unsurprising that AR(16) once again outperforms AR(4), given the setup differs little from that in [Lew22], bar the addition of the extra 2022 Q3 datapoint. More noteable is how AR(16) outperforms the constant growth model, with its hybrid model again lying in between.

The relative success of the AR(16) model is especially surprising given the adage that you would want 30 observations per coefficient you estimate. By this logic, AR(2) should perform best, but as documented in [Lew22] this does not prove to be the case in our experiments.

Finally, we comment briefly on our side quest, the task of predicting economic growth had the Covid-19 pandemic not occurred (see Appendix III). Note first (as observed in [Lew22]) that our AR(4) model predicts continued economic growth, without the 2020 slump. However, by the year 2021 actual growth has caught up with these predictions, and in 2022 significantly exceeds our model's predictions.

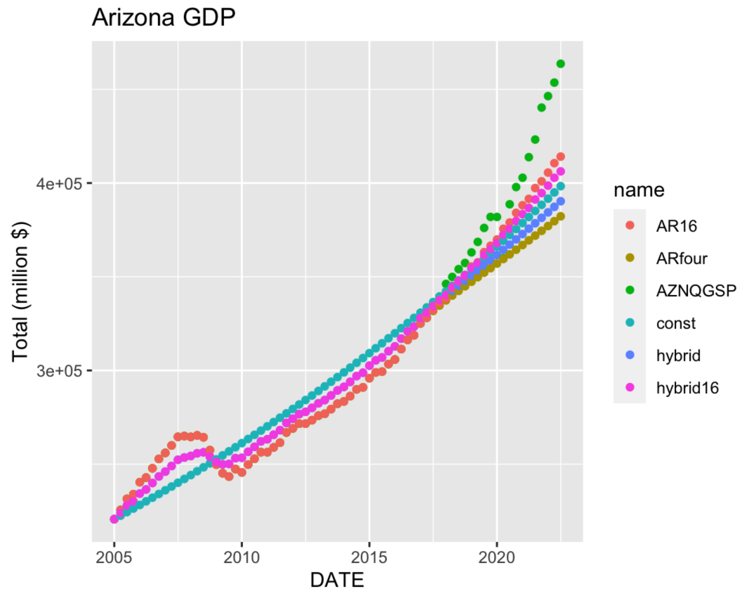

FIGURE 3. Step 2: note how AR(16) outperforms the constant growth model in this case.

Repeating this experiment with different model specifications, we find that the same is true for the constant growth model, only this time the constant growth model performs less well than AR(4). Conversely, the AR(16) model provides a remarkably accurate projection of economic growth under this counterfactual, with MSFE 2.5 times smaller than that of the AR(4) model (see Table 12.1).

9. CONCLUSION

In this project we have applied autoregressive, constant growth and hybrid models to real-world data to predict future economic growth.

Our experimental evidence continues to point to the AR(16) model as a better predictor of economic growth (indicated by lower MSFE) contrary to empirical evidence that an AR(4) model should perform best.

As in [Lew22], we may explain this unexpected result by pointing to the small sample size upon which our predictions are made. A more reliable estimate of MSFE could perhaps be obtained by k-fold cross-validation, but this task falls beyond the scope of this project. So, once again, we have insufficient evidence to reject the hypothesis that AR(4) models provide the best prediction of future economic growth. The task continues to warrant further investigation!

REFERENCES

[DGC+19] Jesse Dodge, Suchin Gururangan, Dallas Card, Roy Schwartz, and Noah A. Smith (2019), `https://arxiv.org/pdf/1909.03004.pdf`. Accessed December 4, 2022. ↑10

[FRE22] U.S. Bureau of Economic Analysis FRED Federal Reserve Bank of St. Louis, *Gross Domestic Product: All Industry Total* (2022), `https://fred.stlouisfed.org/series/ALNQGSP`. ↑3

[Lew22] Daniel Lewis (2022), `https://github.com/danlewis92/GDPproject/blob/main/EconometricsGDPProject%20(3).pdf`. Accessed February 20, 2023. ↑1, 3, 4, 5, 6, 8, 8, 9, 11

[Yul27] G. Udny Yule, *On a method of investigating periodicities disturbed series, with special reference to Wolfer's sunspot numbers*, Phil. Trans. of the Royal Soc. **226** (1927), DOI https://doi.org/10.1098/rsta.1927.0007. ↑7.1

[Wal31] Gilbert Walker, *On periodicity in series of related terms*, Proc. of the Royal Soc. **131** (1931), DOI https://doi.org/10.1098/rspa.1931.0069. ↑7.1

10. Appendix I

We make a few remarks concerning the reproducibility of our results., following the example of [DGC+19].

- Description of computing infrastructure
  All computation was run on the author's Macbook Air with a 1.1 GHz Dual-Core Intel Core i3 processor and 8GB memory.
- Average runtime for each approach
  See Table 8.1 and Table 8.2 in Section 8.
- Details of train/validation/test splits
  See Section 7.
- Corresponding validation performance for each reported test result
  There are not separate validation and test sets in this paper.
- A link to implemented code
  See `https://github.com/danlewis92/GDPproject`.

11. Appendix II

We include here the results of our exploratory data analysis for the state of Arizona. The AR(4) predictions of [Lew22] are plotted in red, while the true GDP 2022 Q3 figure is colored green. Note that real-life GDP growth surpassed our AR(4) prediction, by $463,653.7 - 461,270.2 = \$2,383.5$ million.

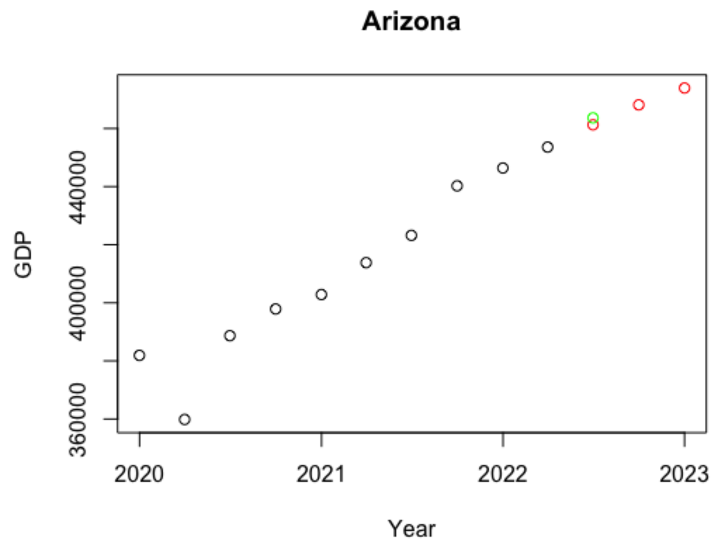

FIGURE 4. Arizona GDP

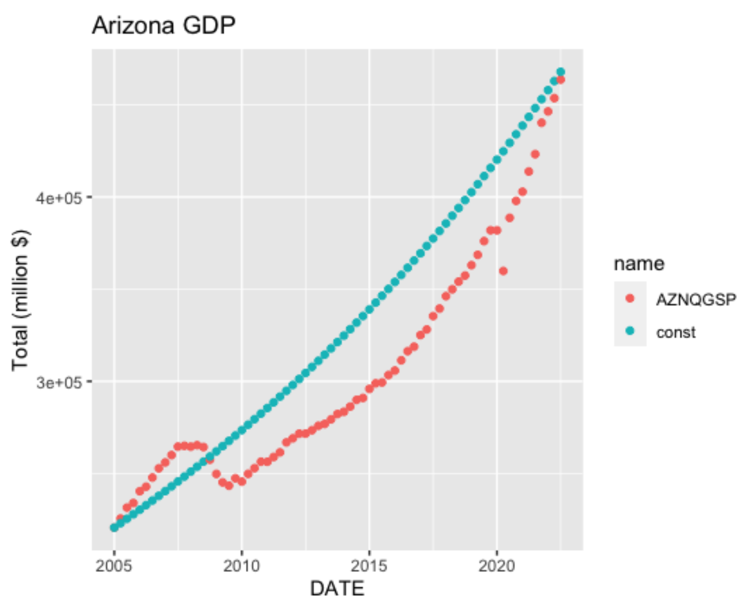

FIGURE 5. Arizona GDP compared to constant growth prediction

The second plot compares the true growth figures, in red, against the straight line, in blue, provided by a constant growth model. In Section 4 we describe how the constant growth rate is computed.

12. Appendix III

We include here the results of our side quest predicting economic growth had the Covid-19 pandemic not occurred.

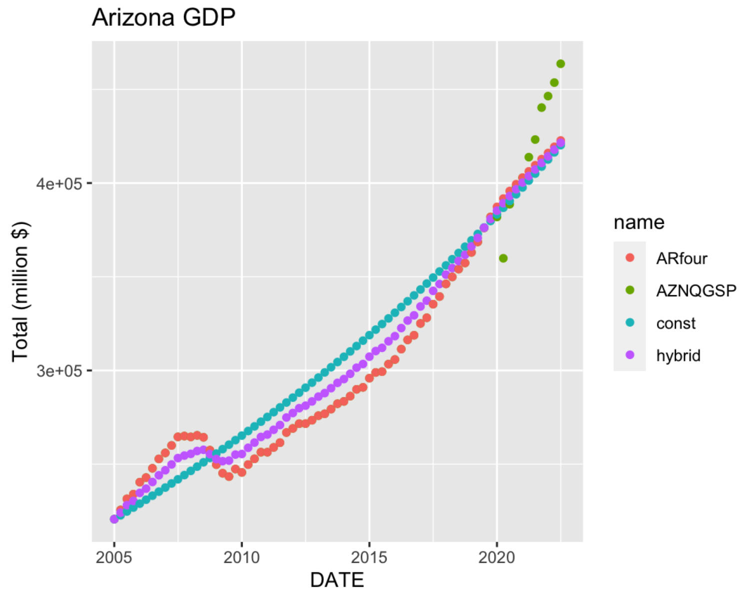

FIGURE 6. Note how the AR(4) model outperforms the constant growth model in this case.

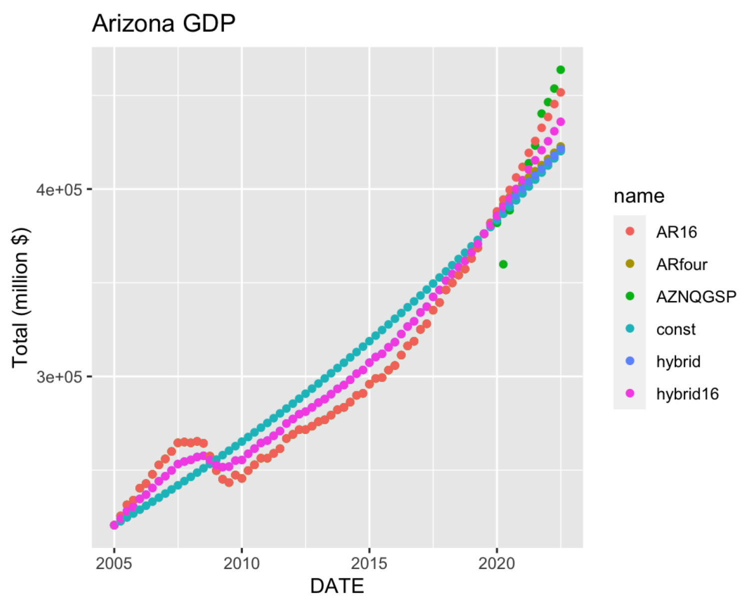

FIGURE 7. Note the remarkable accuracy of the AR(16) model when trained on 60 datapoints to predict the remaining 11 datapoints.

Model	MSFE
AR(4)	425338645
Const.	606407112
Hybrid	502906627
AR(16)	170106418
Hybrid16	289876618

Table 12.1. A counterfactual: we train each model on the first 60 datapoints in our sample and predict the remaining 11 datapoints as if the Covid-19 pandemic had not occurred.