ADVANCED TIME SERIES ANALYSIS ASSIGNMENT

Artur Dragunov

TIME SERIES DEFINITION

- Source: Federal Housing Finance Agency (FHFA)[1]
- Time series: House Price Index (HPI) of Texas for the period 1991Q1:2022Q2;
- Specifics: Not seasonally adjusted (NSA) quarterly time series estimated exclusively with data from purchase-money mortgages (purchase-only).

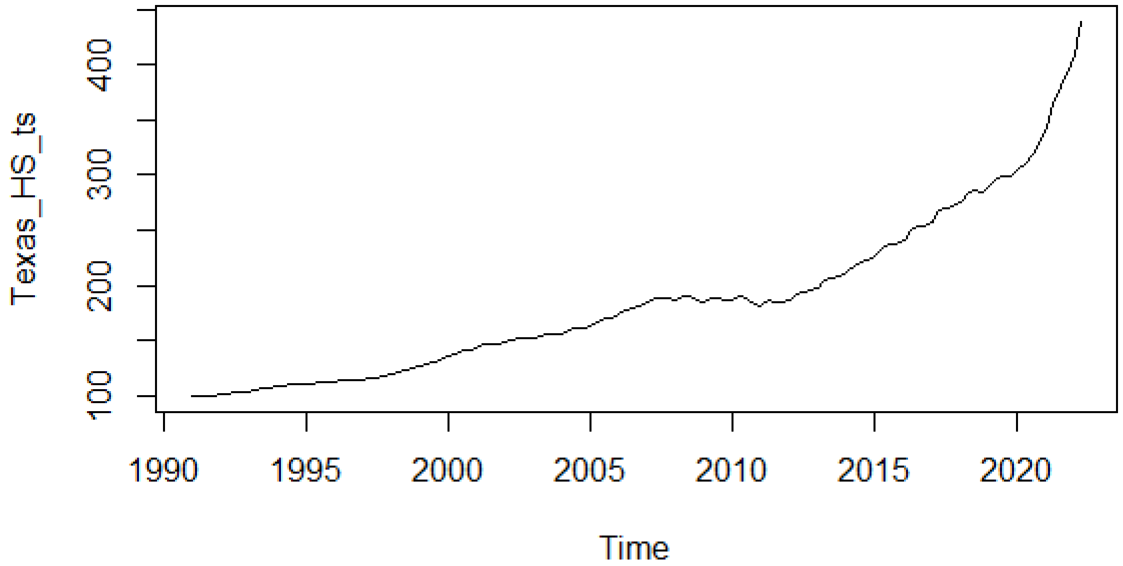


Figure 1. Texas HPI in levels. Source: FHFA

TIME SERIES SPECIFICATION

- Time series in levels is not stationary, it has a positive trend;
- ADF-test with trend showed that the trend is stochastic (H0: trend stochastic):

```
CADFtest(Texas_HS_ts_log, type= "trend", criterion= "BIC", max.lag.y=max.lag)
              ADF test

data:  Texas_HS_ts_log
ADF(9) = 0.73863, p-value = 0.9997
alternative hypothesis: true delta is less than 0
sample estimates:
     delta
0.01025353
```

- Differenced time series has seasonality. Q2 on average has higher index than other quarters, which is expected[2];
- After seasonal differencing time series becomes stationary with 5% significance level (H0: time series is not stationary).

```
 CADFtest(Texas_HS_ts_log_ds, type= "drift", criterion= "BIC",
max.lag.y=max.lag)

              ADF test

data:  Texas_HS_ts_log_ds
ADF(4) = -3.0785, p-value = 0.03115
alternative hypothesis: true delta is less than 0
sample estimates:
     delta
-0.3476359
```

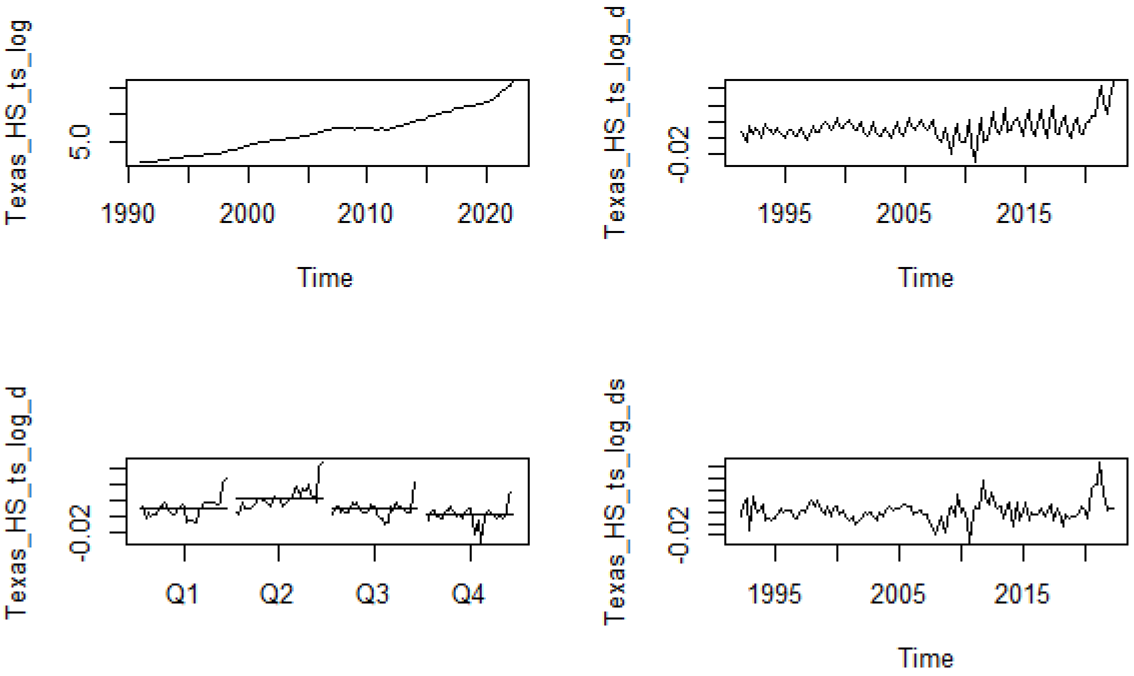


Figure 2. Starting from the left-upper corner: Log(Texas_HS), log(Texas_HS) in first differences, quarterly plots, log(Texas_HS) in first and seasonal differences.

MODEL DEFINITION

- Both ACF and PACF demonstrate that log(Texas_HS) in first and seasonal differences is not a white noise;
- Charts suggest to use SARMA models, because PACF demonstrates significant partial autocorrelations at 1- and 3-year lags. Similar case is with ACF.
- Four models are chosen:
    - SARIMA(2,1,0)(3,1,0): AR(2) repeated 3 times;
    - SARIMA(0,1,3)(0,1,3): MA(3) repeated 3 times;
    - SARIMA(1,1,1)(3,1,3): ARIMA combination of both MA and AR models repeated 3 times;
    - SARIMA(4,1,0)(3,1,0): AR(4) repeated 3 times.

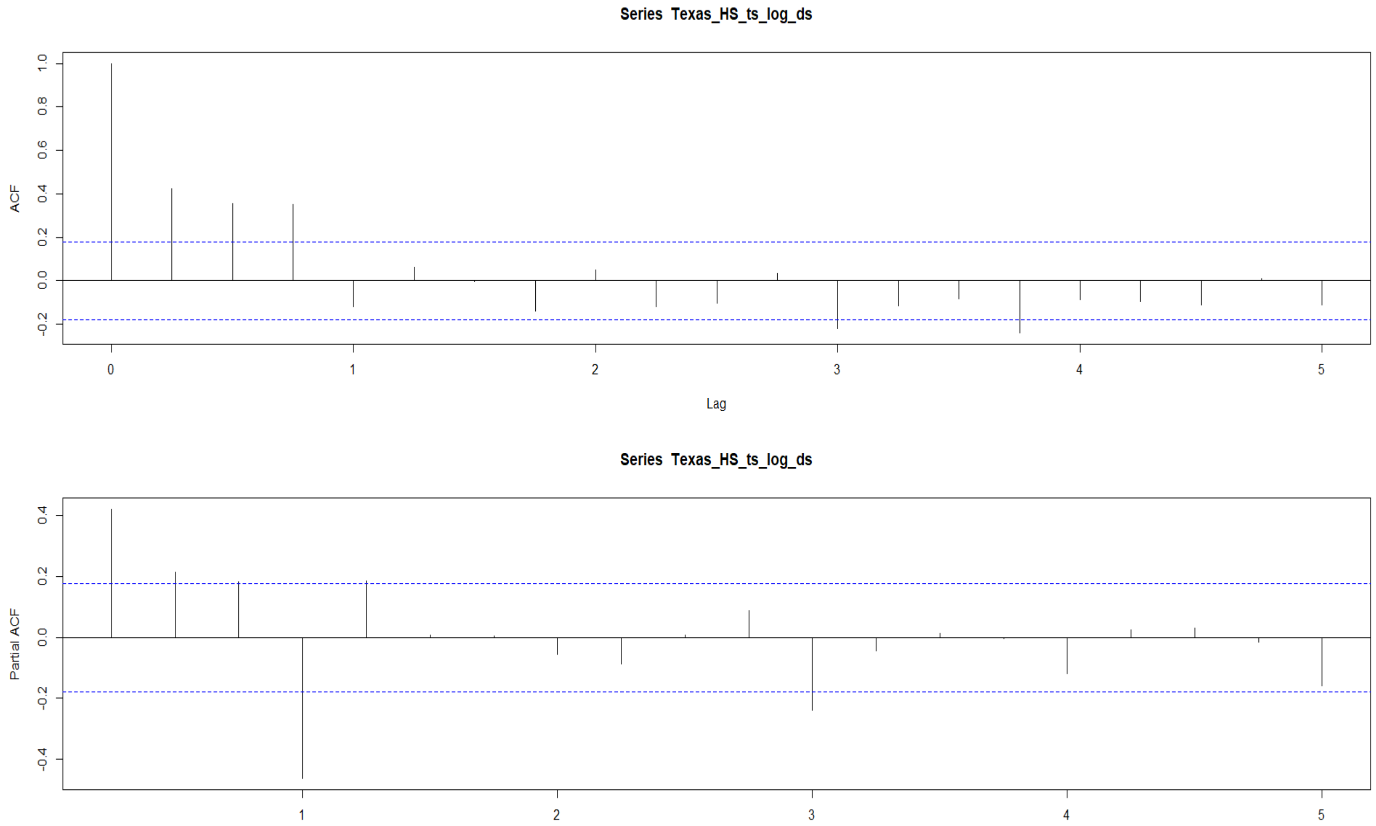


Figure 3. Correlogram and partial correlogram of log(Texas_HS) in first and seasonal differences.

MODEL VALIDATION

All four models are validated based on residual correlograms and Ljung-Box tests (H0: residuals are white noise).

Example of the test script and output:

```
Box.test(fit_sarma1$res, lag=10, type="Ljung-Box")

              Box-Ljung test

data:  fit_sarma1$res
X-squared = 8.4618, df = 10, p-value = 0.5838
```

Lag 10 is sufficient as we have quarterly data.

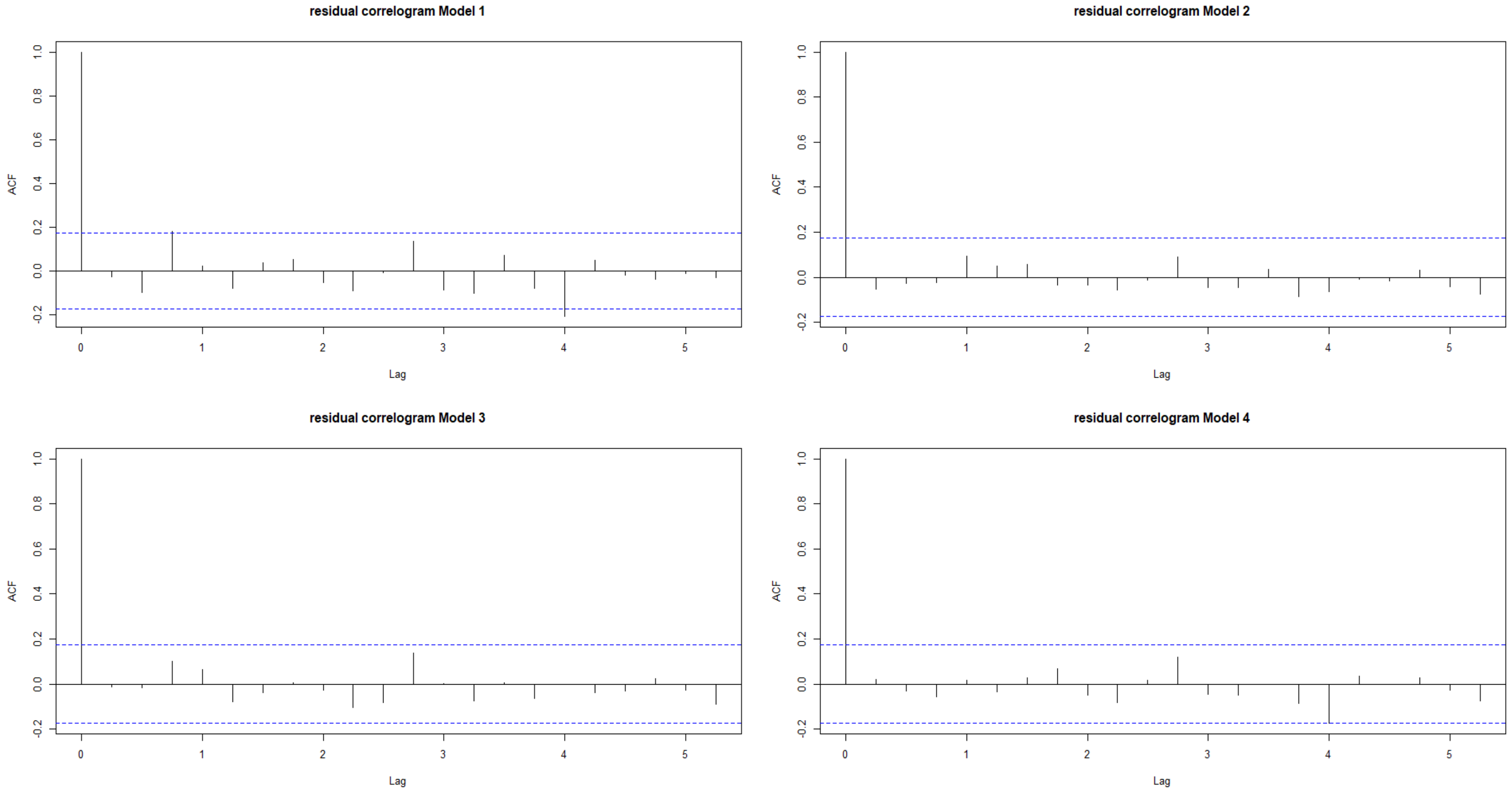


Figure 4. Residual correlograms of the SARIMA models

FORECASTING AND MODEL COMPARISON

- All four models have similar forecasts (red line) and prediction intervals (PI) (blue lines), see Figure 5;
- PI become wider with increasing h, as uncertainty increases;
- All forecasts show that the positive trend will remain but in 2023 it will be slower – Texas real estate prices should keep slowly growing based on the forecasts.

Table 1 – Table of performance:

- Model 2 has the lowest AIC and BIC, but the highest Mean Absolute Error (MAE);
- Model 3 doesn't have all the highest order terms significant;
- Model 1 and Model 4 have in practice good correlograms, almost the same BIC and their forecasting performance is not statistically different. BUT Model 1 has the worst AIC. Thus, out of these four models, we would pick **Model 4** as a best model.

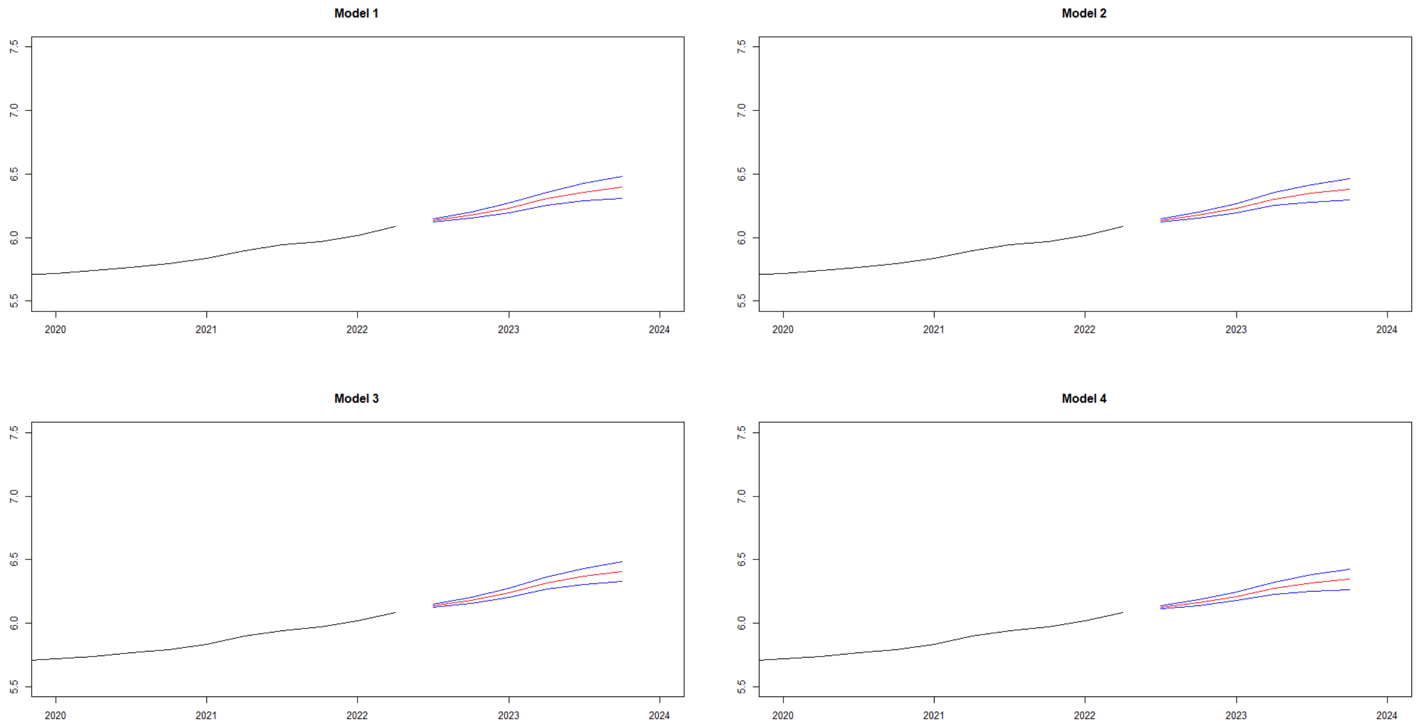


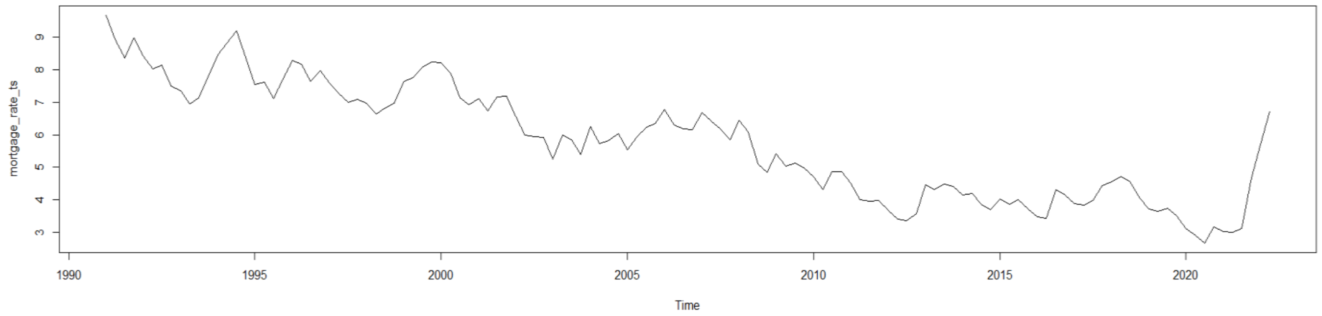
Figure 5. 6-step-ahead forecast of log(Texas_HS) using Models 1-4

Model	Highest order	Correlogram	Box Test	AIC	BIC	MAE	# param
Model 1	yes	ok in practice	white noise	-849,94	-832,92	0,0067	5
Model 2	yes	ok	white noise	-860,25	-840,40	0,0098	6
Model 3	no	ok	white noise	-858,40	-832,87	0,0059	8
Model 4	yes	ok in practice	white noise	-855,43	-832,73	0,0070	7

Table 1. Table of performance of four SARIMA models

MULTIVARIATE TIME SERIES ANALYSIS

- Source: Federal Reserve Economic Data (FRED)[3], [4]
- Time series: 30-year fixed rate mortgage average in the US; M2 - money stock measure
- Specifics: Both are quarterly (end of the period) for 1991Q1-2022Q2. Mortgage rate are not seasonally adjusted percentage points. M2 - as a proxy for Money Supply in the US - seasonally adjusted billions of dollars.
- Both time series have a stochastic trend and are I(1) (checked analogously to the slide 3).
- A research hypothesis: Growth of M2 should have a positive relationship and mortgage rates a negative relationship with the growth of the Texas HPI.



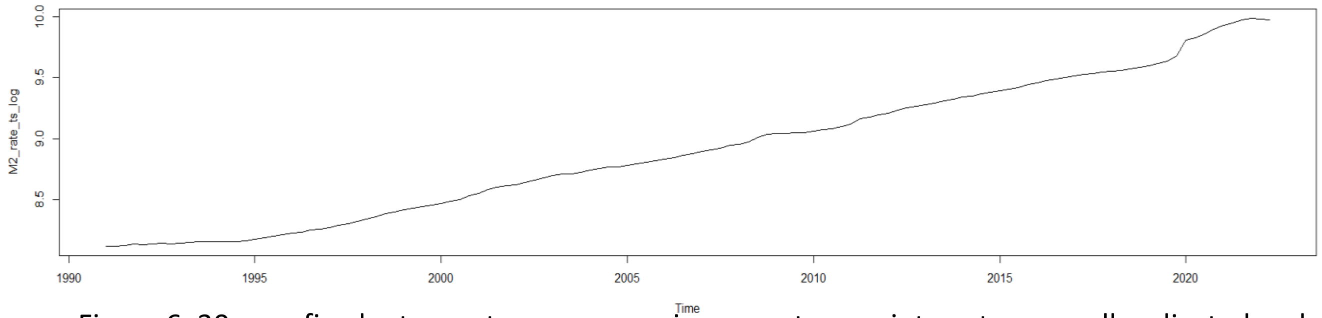


Figure 6. 30-year fixed rate mortgage average in percentage points not seasonally adjusted and log(M2) seasonally adjusted. Source: FRED

VECTOR ERROR CORRECTION MODEL

- The order of the VAR model in levels based on SIC is 6.

```
    VARselect(df_ts,lag.max=10,type="const")
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
    10      6      6     10
```

- Test for cointegration using Johansen's trace test statistic – we have two cointegration relations, see Table 2:

H0: r=0 – we rejected as 52.02 > 34.91

H0: r=1 – we rejected as 21.19>19.96

H0: r=2 – we didn't reject as 1.61 < 9.24

- Three equations are estimated by OLS, see Table 3. Not all ECT are negative, which is against the theoretical expectations. BUT some of the positive estimators are not statistically significant, which would imply that the signs are not essential;

- Estimated cointegration vectors:
    - Log(M2_rate) =-0.29*Mortgage_rate +10.06
    - Log(Texas_HS) =-0.17*Mortgage_rate+5.63.

H0:	test	10pct	5pct	1pct
r <= 2	1.61	7.52	9.24	12.97
r <= 1	21.19	17.85	19.96	24.60
r = 0	52.02	32.00	34.91	41.07

Table 2. Part of the output from the Johansen's trace test

```
Call:
lm(formula = substitute(form1), data = data.mat)

Coefficients:
                      M2_rate_ts_log.d  Texas_HS_ts_log.d  mortgage_rate_ts.d
ect1                  -2.825e-02          2.501e-02         -4.053e+00
ect2                   5.130e-02         -3.465e-02          3.460e+00
M2_rate_ts_log.dl1     2.877e-01         -9.723e-02         -1.675e-01
Texas_HS_ts_log.dl1    3.745e-02          4.259e-01         -3.742e-01
mortgage_rate_ts.dl1  -8.215e-03          1.406e-03          3.456e-01
M2_rate_ts_log.dl2     1.240e-01          1.424e-01         -6.199e-01
Texas_HS_ts_log.dl2   -1.511e-01         -2.205e-01          1.046e+01
mortgage_rate_ts.dl2   2.082e-03         -4.608e-03          1.611e-01
M2_rate_ts_log.dl3     5.532e-02          5.386e-02          5.173e+00
Texas_HS_ts_log.dl3    1.126e-01          2.540e-01          6.472e+00
mortgage_rate_ts.dl3   1.277e-06          2.137e-03          1.617e-01
M2_rate_ts_log.dl4     2.685e-02          4.170e-02          2.006e-01
Texas_HS_ts_log.dl4   -1.323e-01          6.487e-01          6.723e+00
mortgage_rate_ts.dl4  -4.250e-03         -2.864e-03         -9.232e-03
M2_rate_ts_log.dl5    -2.362e-02          2.512e-01         -2.798e+00
Texas_HS_ts_log.dl5   -1.681e-01         -1.798e-01          6.655e+00
mortgage_rate_ts.dl5  -1.565e-03          6.178e-04         -2.235e-02


$beta
                           ect1       ect2
M2_rate_ts_log.l1     1.0000000  0.0000000
Texas_HS_ts_log.l1    0.0000000  1.0000000
mortgage_rate_ts.l1   0.2886829  0.1664343
constant            -10.0575833 -5.6299552
```

Table 3. Output of the Vector Error Correction Model

VECTOR ERROR CORRECTION MODEL

- Table 4 reports coefficients from Table 3 which are significant under 5% significance level.
- ECT of log(M2_rate).d and log(Texas_HS).d are not significant – time series are not responsive to deviations from equilibrium.
- Both ECT1 and ECT2 for mortgage_rate.d are significant, big, and have positive and negative signs. The reason of instability might be due to the recent increases of interest rates by Federal Reserve Board (FED)[5] – actions heavily against the trend.
- log(M2_rate).d coefficients in log(Texas_HS).d equation > 0. Increase of M2 growth corresponded with increase of Texas_HS growth. Mortgage_rate.d has a negative sign, as was expected.
- 89% of Texas_HS_ts_log.d variance can be explained by the regressors from Table 3, see R-squared.
- All three models are jointly significant.

	M2_rate_ts_log.d
M2_rate_ts_log.dl1	2.877e-01
mortgage_rate_ts.dl1	-8.215e-03
R-squared	0.7025
model p-value	< 2.2e-16
	Texas_HS_ts_log.d
Texas_HS_ts_log.dl1	0.4259015
M2_rate_ts_log.dl2	0.1424363
Texas_HS_ts_log.dl2	-0.2205317
mortgage_rate_ts.dl2	-0.0046081
Texas_HS_ts_log.dl3	0.2540491
Texas_HS_ts_log.dl4	0.6487438
M2_rate_ts_log.dl5	0.2512357
R-squared:	0.8914
model p-value	< 2.2e-16
	mortgage_rate_ts.d
ect1	-4.052643
ect2	3.459674
mortgage_rate_ts.dl1	0.345556
Texas_HS_ts_log.dl2	10.463664
R-squared	0.3834
model p-value	1.315e-05

Table 4. Significant estimators from Table 3, R-squared and joint p-value.

FORECASTING

- VECM model transformed to VAR specification in levels. VAR model was used to forecast
- Based on the forecast, money supply will continue smoothly growing for the next six quarters, see Figure 7;
- Log(Texas_HS) is forecasted to continue growing slowly till the end of the year but then lie on a plateau in 2023. The forecast is similar to SARIMA models but might be more realistic regarding plateau. The sentiments on the market are negative[6],[7];
- It is forecasted that the mortgage rates will decrease and move back to their downtrend. The forecast goes against the expectations of the FED, which intends to keep increasing the interest rates at least till the end of 2023[5].

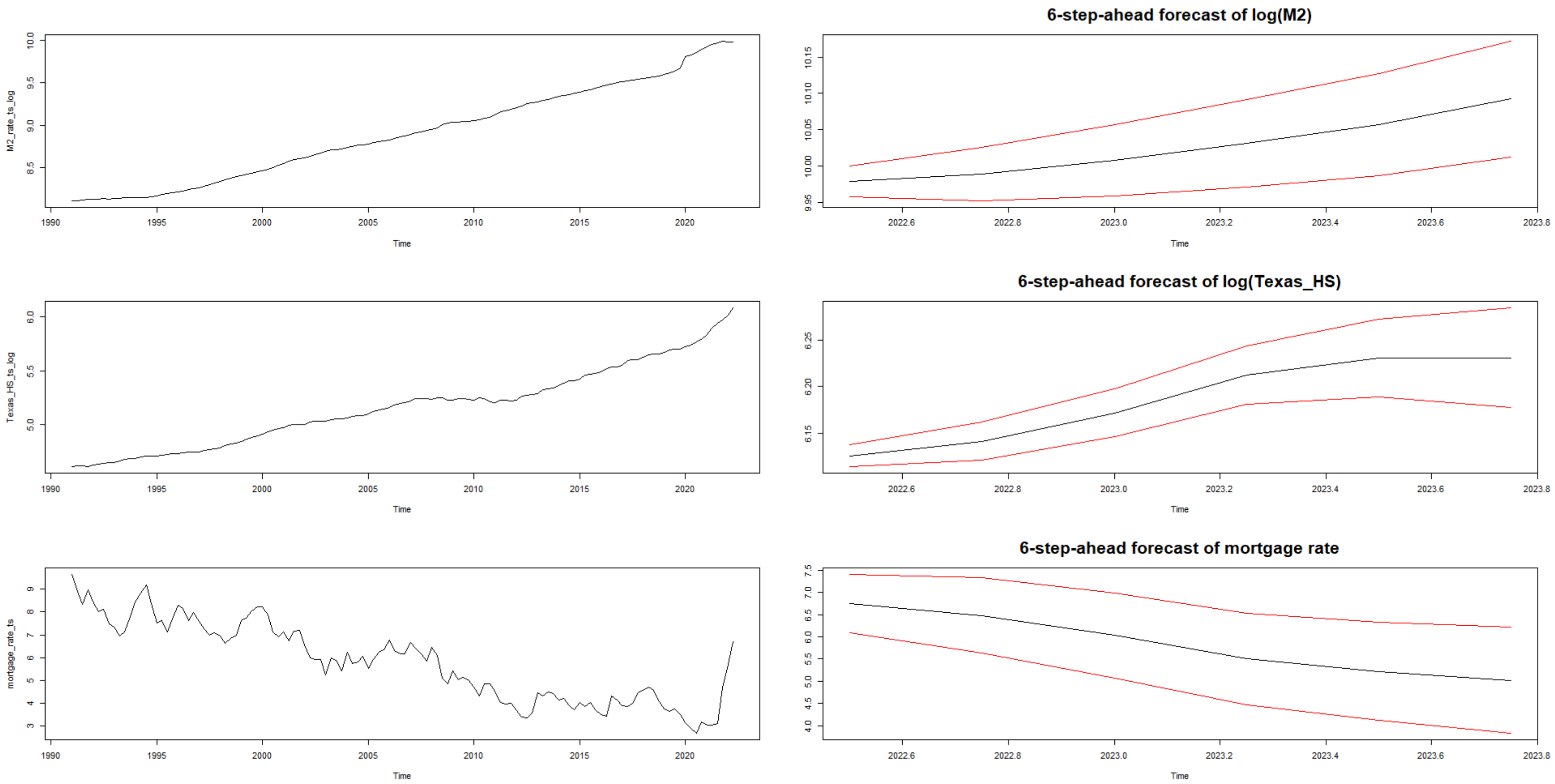


Figure 7. Log(M2), log(Texas_HS) and mortgage rates to the left and their 6-step-ahead forecasts together with 95% prediction interval to the right

CONCLUSION

- HPI for Texas has a strong seasonal correlation even after first and seasonal differencing;
- SARIMA models were implemented to describe the behaviour of the time series;
- Model 4: SARIMA(4,1,0)(3,1,0) showed the best performance;
- Log(M2) and mortgage rates were used for the multivariate analysis. Both time series were not stationary with stochastic trends;
- Johansen's trace test showed two cointegrations, thus a VECM model was used;
- VECM output indicates a positive effect of M2 growth and a negative of mortgage rates on Texas HPI.
- 6-step-ahead forecast for log(M2) shows that the money supply growth will continue. Forecast of log(Texas House Price Index) looks similar to the one achieved with SARIMA models. Forecast of the mortgage rates shows decrease of the percentage points for the next six quarters.

SOURCES OF INFORMATION

1) https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index.aspx
2) https://www.realtor.com/research/september-2022-data/
3) https://fred.stlouisfed.org/series/M2SL#0
4) https://fred.stlouisfed.org/series/MORTGAGE30US
5) https://www.bankrate.com/banking/federal-reserve/when-will-the-fed-stop-raising-rates/
6) https://www.forbes.com/sites/qai/2022/11/26/real-estate-trends-homebuilder-sentiment-drops-along-with-housing-prices/?sh=7ff8ab266958
7) https://www.marketplace.org/2022/11/25/housing-market-less-appealing-to-consumers/