

# ARMA-X Analysis Tutorial

## Contents

<b>S&amp;P500 Univariate ARMA-X Models</b>	<b>2</b>
Tweet Dummy as Exogenous . . . . .	2
Tweet Count as Exogenous . . . . .	7
Tariff as Exogenous . . . . .	13
Trade Mention as Exogenous . . . . .	19
China Mention as Exogenous . . . . .	25
Positive Vibe as Exogenous . . . . .	31
Negative Vibe as Exogenous . . . . .	37

	Model 1
ar1	0.9828*** (0.0017)
ma1	-0.6786*** (0.0073)
ma2	-0.2118*** (0.0087)
ma3	-0.0120 (0.0080)
ma4	0.0331*** (0.0071)
intercept	0.0202*** (0.0041)
dummy_lag_0	0.0013*** (0.0002)
dummy_lag_1	0.0007*** (0.0002)
dummy_lag_2	0.0001 (0.0002)
AIC	-45719.7236
AICc	-45719.7126
BIC	-45640.7043
Log Likelihood	22869.8618
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

Table 1: ARMAX Model Results

## S&P500 Univariate ARMA-X Models

### Tweet Dummy as Exogenous

```
#auto.armax selects the lowest AIC value given r (exogenous variable lags)
res1 = auto.armax(data$SPY_vol, xreg=data$dummy, nb.lags=2,
                  latex=T, max.p = 6, max.q = 6, max.d=0)
```

```
#armax enables a custom armax specification with p,q,r
res2 = armax(data$SPY_vol, xreg=data$dummy, nb.lags=2,
             p=5, q=0, d=0, latex=T)
```

```
#auto.armax.r selects the lowest AIC checking all 3 p,q,r values
res3 = auto.armax.r(data$SPY_vol, x=data$dummy,
                    max_p = 6, max_q = 6, max_r = 6, criterion = "AIC", latex=T)
```

```
#we want to plot the IRFs of these models
nb.periods = 7

irf.plot(res1, nb.periods)
```

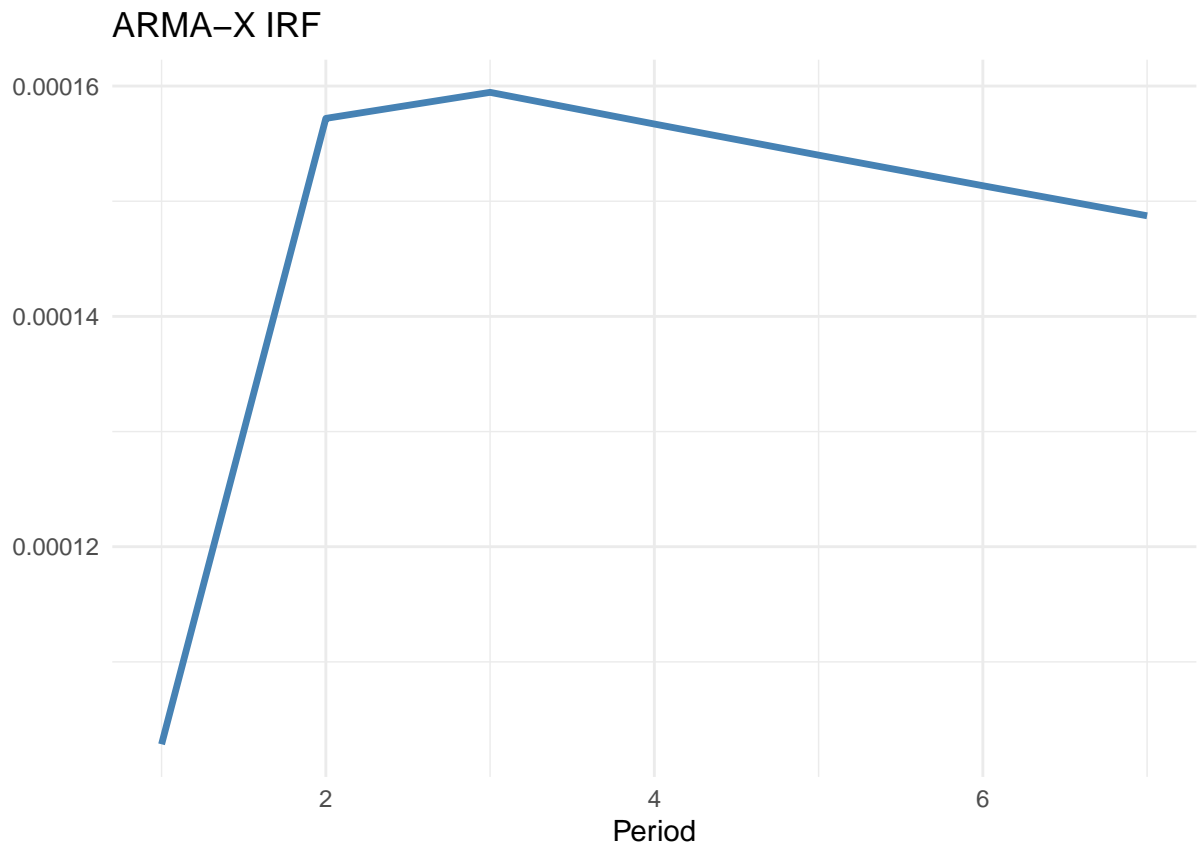
	Model 1
ar1	0.3576*** (0.0071)
ar2	0.0416*** (0.0075)
ar3	0.0994*** (0.0074)
ar4	0.1045*** (0.0075)
ar5	0.0816*** (0.0071)
intercept	0.0199*** (0.0018)
dummy_lag_0	0.0015*** (0.0002)
dummy_lag_1	0.0009*** (0.0002)
dummy_lag_2	0.0001 (0.0002)
AIC	−44706.1942
AICc	−44706.1832
BIC	−44627.1749
Log Likelihood	22363.0971
Num. obs.	19969
*** $p < 0.001$ ; ** $p < 0.01$ ; * $p < 0.05$	

Table 2: ARMAX Model Results

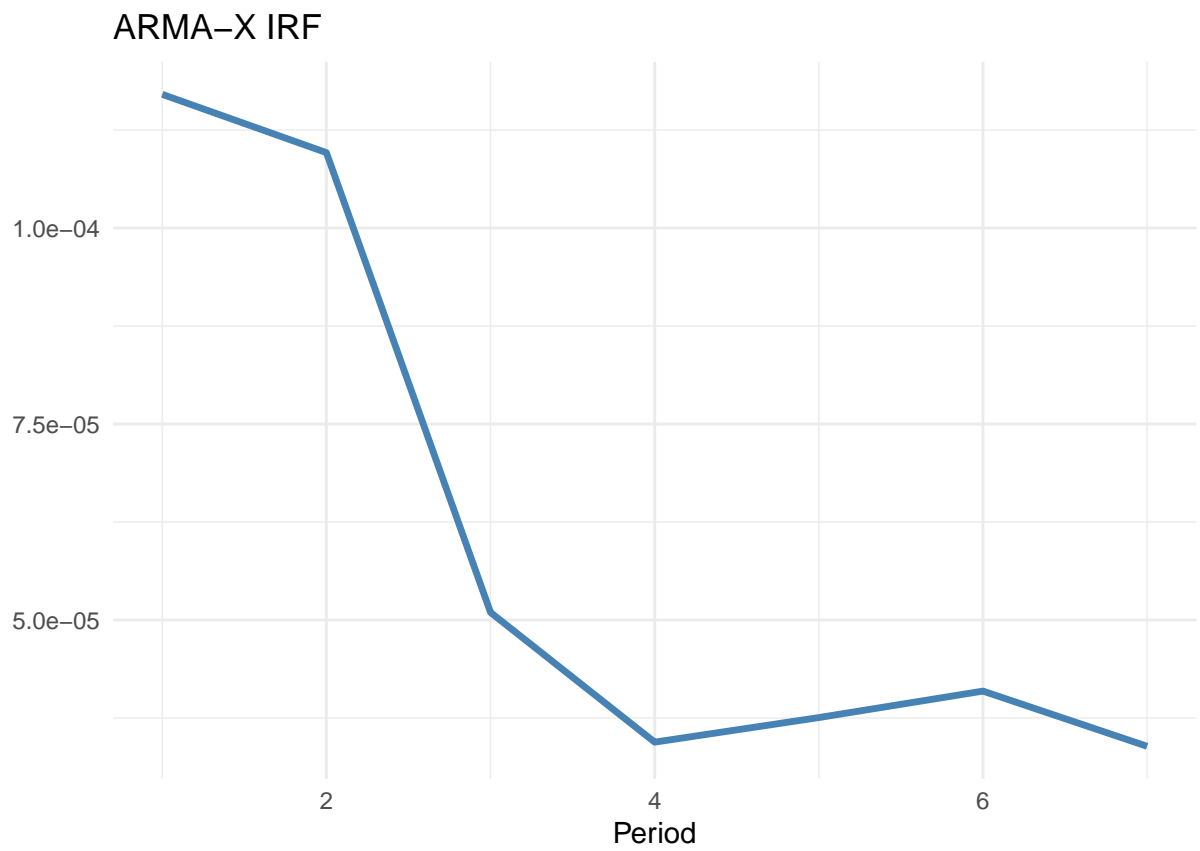
	Model 1
ar1	−0.6515*** (0.0104)
ar2	0.0295* (0.0133)
ar3	0.0193* (0.0085)
ar4	0.1384*** (0.0122)
ar5	0.6404*** (0.0111)
ar6	0.6864*** (0.0075)
ma1	0.9508*** (0.0078)
ma2	0.2802*** (0.0137)
ma3	0.1832*** (0.0117)
ma4	0.0692*** (0.0101)
ma5	−0.6142*** (0.0107)
ma6	−0.8002*** (0.0056)
intercept	0.0231*** (0.0042)
dummy_lag_0	0.0010*** (0.0002)
dummy_lag_1	0.0003 (0.0002)
dummy_lag_2	−0.0004 (0.0002)
dummy_lag_3	−0.0006** (0.0002)
dummy_lag_4	−0.0004 (0.0002)
dummy_lag_5	−0.0005** (0.0002)
AIC	−47081.1839
AICc	−47081.1418
BIC	−46923.1482
Log Likelihood	23560.5920
Num. obs.	19966

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

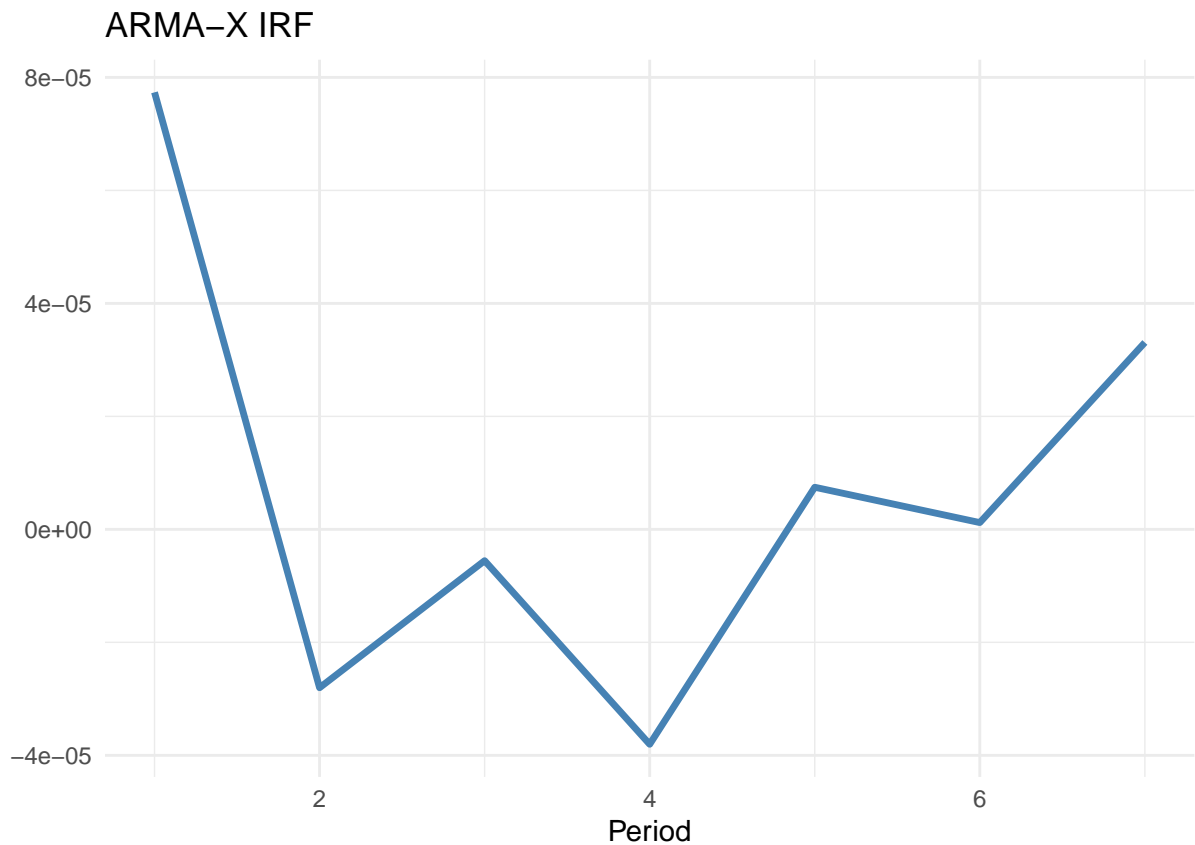
Table 3: ARMAX selected by AIC



```
irf.plot(res2,nb.periods)
```



```
irf.plot(res3$model,nb.periods)
```



## Tweet Count as Exogenous

```
#auto.armax selects the lowest AIC value given r (exogenous variable lags)
res1 = auto.armax(data$SPY_vol, xreg=data$N, nb.lags=2,
                  latex=T, max.p = 6, max.q = 6, max.d=0)
```

```
#armax enables a custom armax specification with p,q,r
res2 = armax(data$SPY_vol, xreg=data$N, nb.lags=2,
             p=5, q=0, d=0, latex=T)
```

```
#auto.armax.r selects the lowest AIC checking all 3 p,q,r values
res3 = auto.armax.r(data$SPY_vol, x=data$N,
                    max_p = 6, max_q = 6, max_r = 6, criterion = "AIC", latex=T)
```

```
#we want to plot the IRFs of these models
nb.periods = 7

irf.plot(res1, nb.periods)
```

	Model 1
ar1	0.9828*** (0.0017)
ma1	−0.6780*** (0.0073)
ma2	−0.2129*** (0.0086)
ma3	−0.0118 (0.0080)
ma4	0.0335*** (0.0071)
intercept	0.0210*** (0.0041)
N_lag_0	0.0003*** (0.0001)
N_lag_1	0.0002** (0.0001)
N_lag_2	0.0000 (0.0001)
AIC	−45696.1228
AICc	−45696.1118
BIC	−45617.1034
Log Likelihood	22858.0614
Num. obs.	19969
*** $p < 0.001$ ; ** $p < 0.01$ ; * $p < 0.05$	

Table 4: ARMAX Model Results



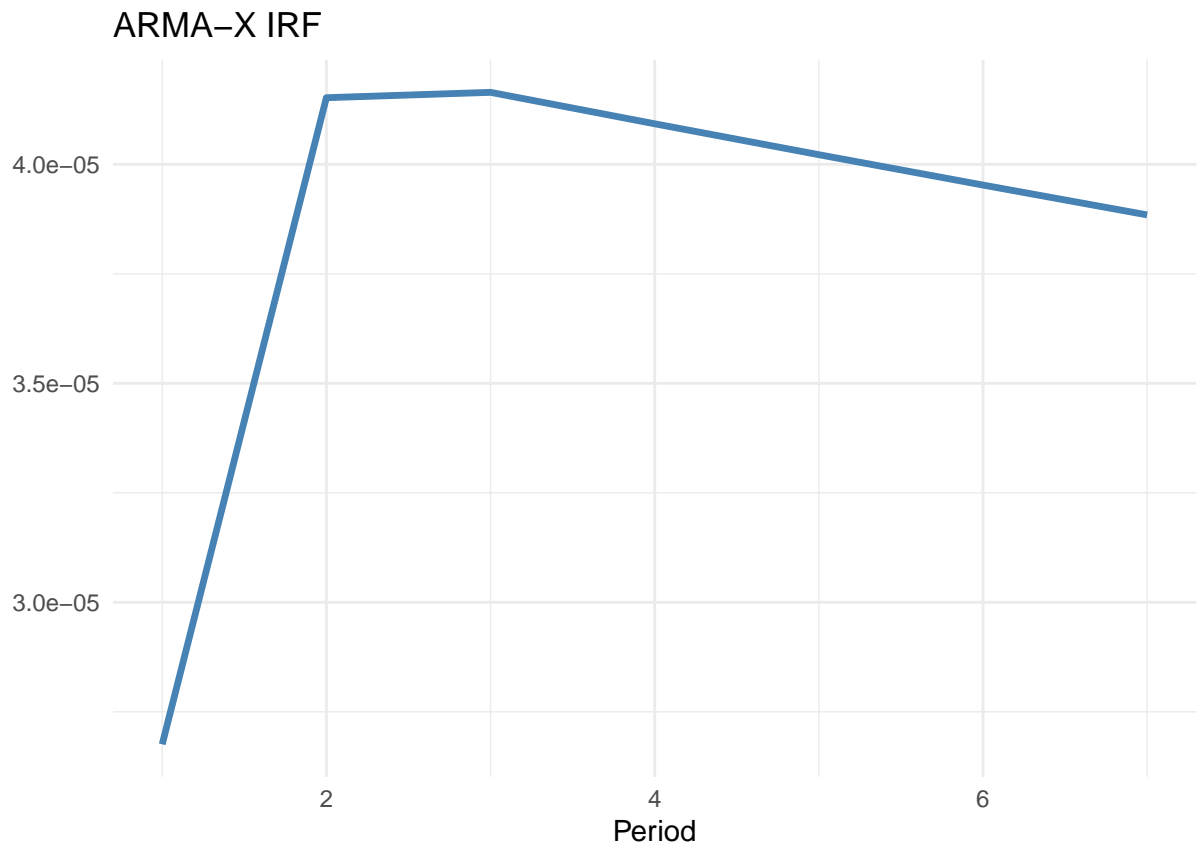
	Model 1
ar1	0.3584*** (0.0071)
ar2	0.0410*** (0.0075)
ar3	0.0991*** (0.0074)
ar4	0.1040*** (0.0075)
ar5	0.0815*** (0.0071)
intercept	0.0208*** (0.0018)
N_lag_0	0.0004*** (0.0001)
N_lag_1	0.0002*** (0.0001)
N_lag_2	0.0000 (0.0001)
AIC	−44677.6875
AICc	−44677.6765
BIC	−44598.6682
Log Likelihood	22348.8438
Num. obs.	19969
*** $p < 0.001$ ; ** $p < 0.01$ ; * $p < 0.05$	

Table 5: ARMAX Model Results

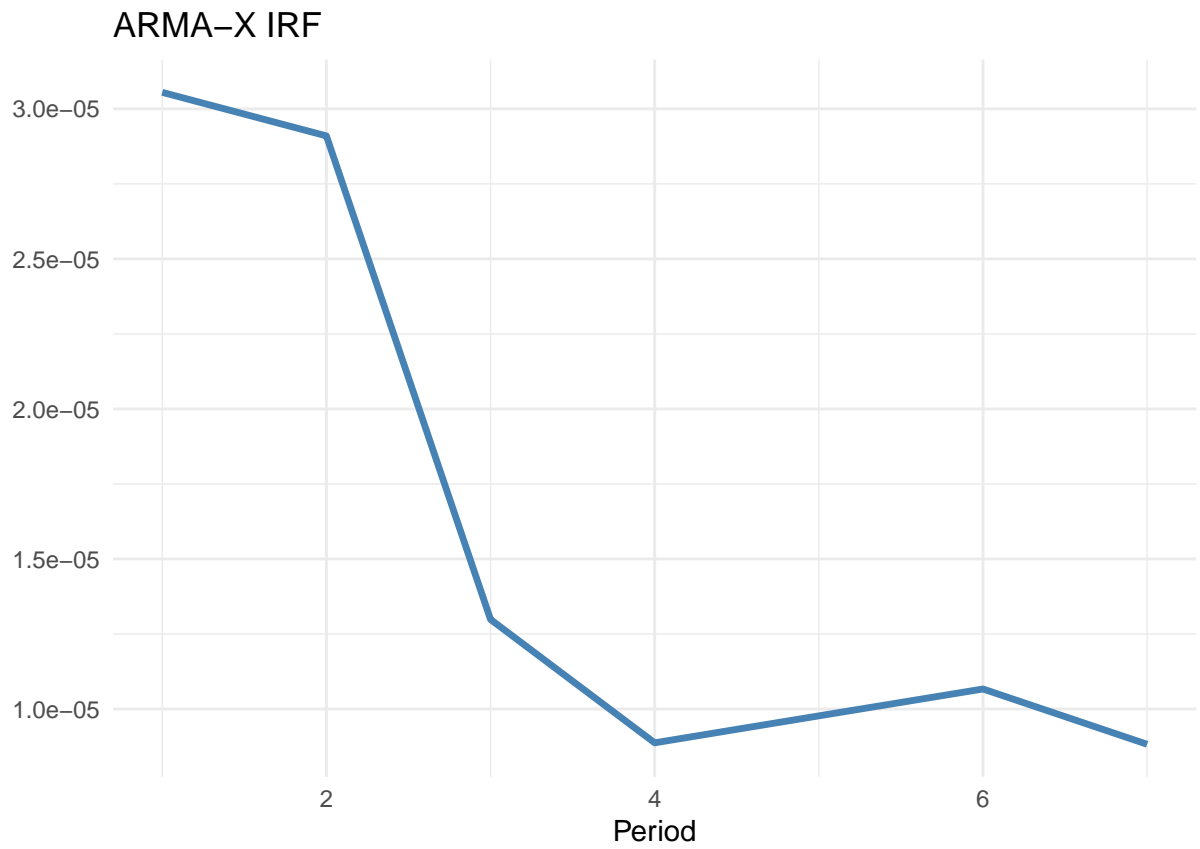
	Model 1
ar1	−0.6509*** (0.0101)
ar2	0.0294* (0.0129)
ar3	0.0195* (0.0085)
ar4	0.1390*** (0.0117)
ar5	0.6406*** (0.0109)
ar6	0.6860*** (0.0075)
ma1	0.9504*** (0.0076)
ma2	0.2801*** (0.0132)
ma3	0.1829*** (0.0115)
ma4	0.0678*** (0.0099)
ma5	−0.6162*** (0.0104)
ma6	−0.8012*** (0.0055)
intercept	0.0213*** (0.0041)
N_lag_0	0.0003*** (0.0001)
N_lag_1	0.0002** (0.0001)
AIC	−47069.4226
AICc	−47069.3953
BIC	−46942.9908
Log Likelihood	23550.7113
Num. obs.	19970

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

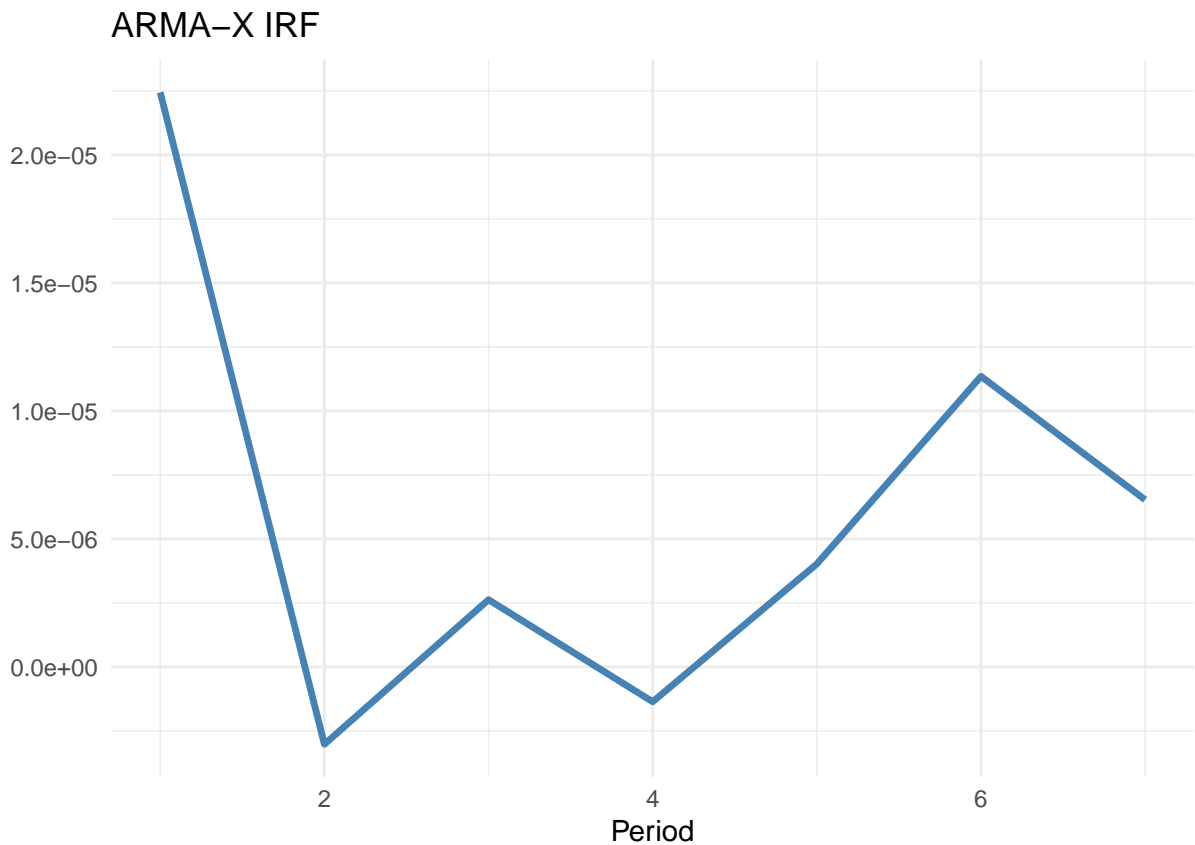
Table 6: ARMAX selected by AIC



```
irf.plot(res2,nb.periods)
```



```
irf.plot(res3$model,nb.periods)
```



## Tariff as Exogenous

```
#auto.armax selects the lowest AIC value given r (exogenous variable lags)
res1 = auto.armax(data$SPY_vol, xreg=data$tariff, nb.lags=2,
                  latex=T, max.p = 6, max.q = 6, max.d=0)
```

```
#armax enables a custom armax specification with p,q,r
res2 = armax(data$SPY_vol, xreg=data$tariff, nb.lags=2,
             p=5, q=0, d=0, latex=T)
```

```
#auto.armax.r selects the lowest AIC checking all 3 p,q,r values
res3 = auto.armax.r(data$SPY_vol, x=data$tariff,
                   max_p = 6, max_q = 6, max_r = 6, criterion = "AIC", latex=T)
```

```
#we want to plot the IRFs of these models
nb.periods = 7

irf.plot(res1, nb.periods)
```

	Model 1
ar1	1.7000*** (0.1313)
ar2	−0.8772*** (0.1518)
ar3	0.1689*** (0.0232)
ma1	−1.3999*** (0.1327)
ma2	0.4605*** (0.1192)
intercept	0.0217*** (0.0040)
tariff_lag_0	0.0042** (0.0014)
tariff_lag_1	0.0199*** (0.0015)
tariff_lag_2	0.0112*** (0.0014)
AIC	−45860.5245
AICc	−45860.5134
BIC	−45781.5051
Log Likelihood	22940.2622
Num. obs.	19969
*** $p < 0.001$ ; ** $p < 0.01$ ; * $p < 0.05$	

Table 7: ARMAX Model Results

	Model 1
ar1	0.3572*** (0.0071)
ar2	0.0427*** (0.0075)
ar3	0.0903*** (0.0075)
ar4	0.0978*** (0.0075)
ar5	0.0859*** (0.0071)
intercept	0.0217*** (0.0017)
tariff_lag_0	0.0047** (0.0015)
tariff_lag_1	0.0201*** (0.0015)
tariff_lag_2	0.0109*** (0.0015)
AIC	−44818.4470
AICc	−44818.4359
BIC	−44739.4276
Log Likelihood	22419.2235
Num. obs.	19969
*** $p < 0.001$ ; ** $p < 0.01$ ; * $p < 0.05$	

Table 8: ARMAX Model Results

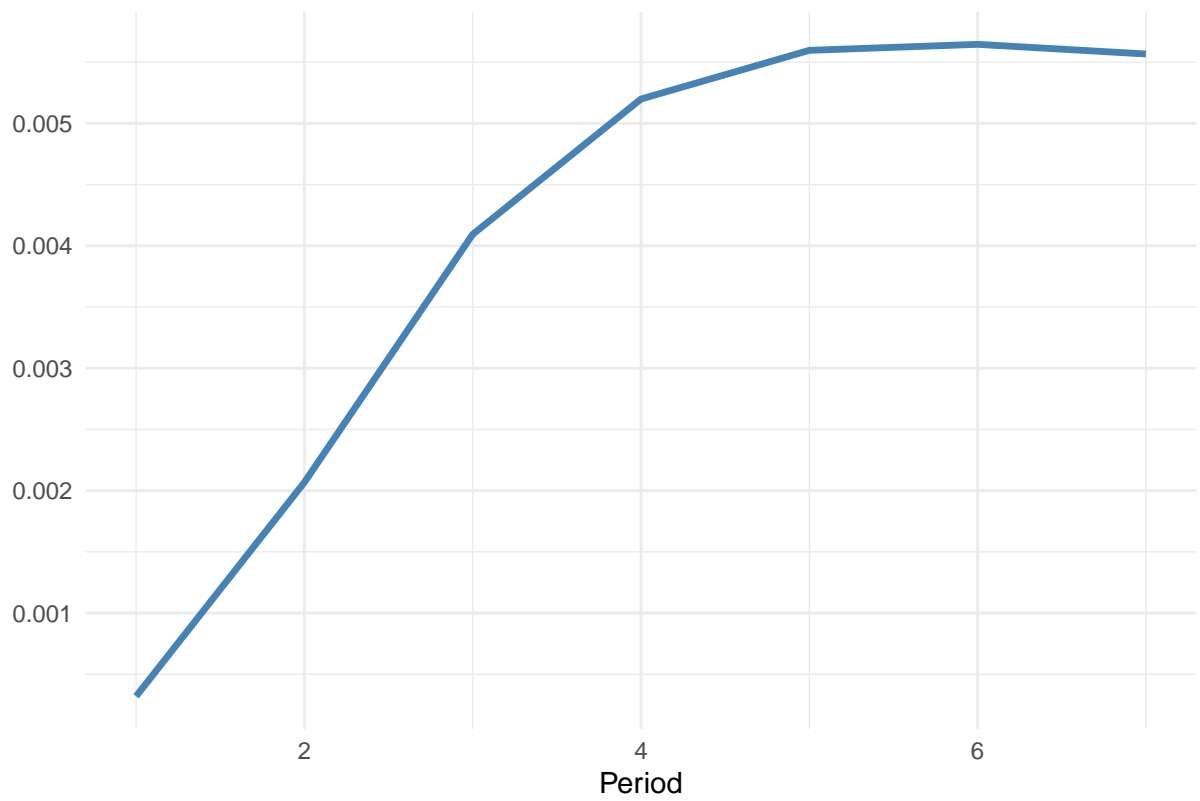
	Model 1
ar1	−0.6474*** (0.0099)
ar2	0.0332** (0.0128)
ar3	0.0164 (0.0089)
ar4	0.1330*** (0.0116)
ar5	0.6417*** (0.0113)
ar6	0.6856*** (0.0076)
ma1	0.9487*** (0.0075)
ma2	0.2737*** (0.0130)
ma3	0.1743*** (0.0118)
ma4	0.0618*** (0.0097)
ma5	−0.6178*** (0.0107)
ma6	−0.7993*** (0.0058)
intercept	0.0217*** (0.0040)
tariff_lag_0	0.0077*** (0.0013)
tariff_lag_1	0.0175*** (0.0014)
tariff_lag_2	0.0093*** (0.0013)
AIC	−47180.9824
AICc	−47180.9517
BIC	−47046.6495
Log Likelihood	23607.4912
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

Table 9: ARMAX selected by AIC

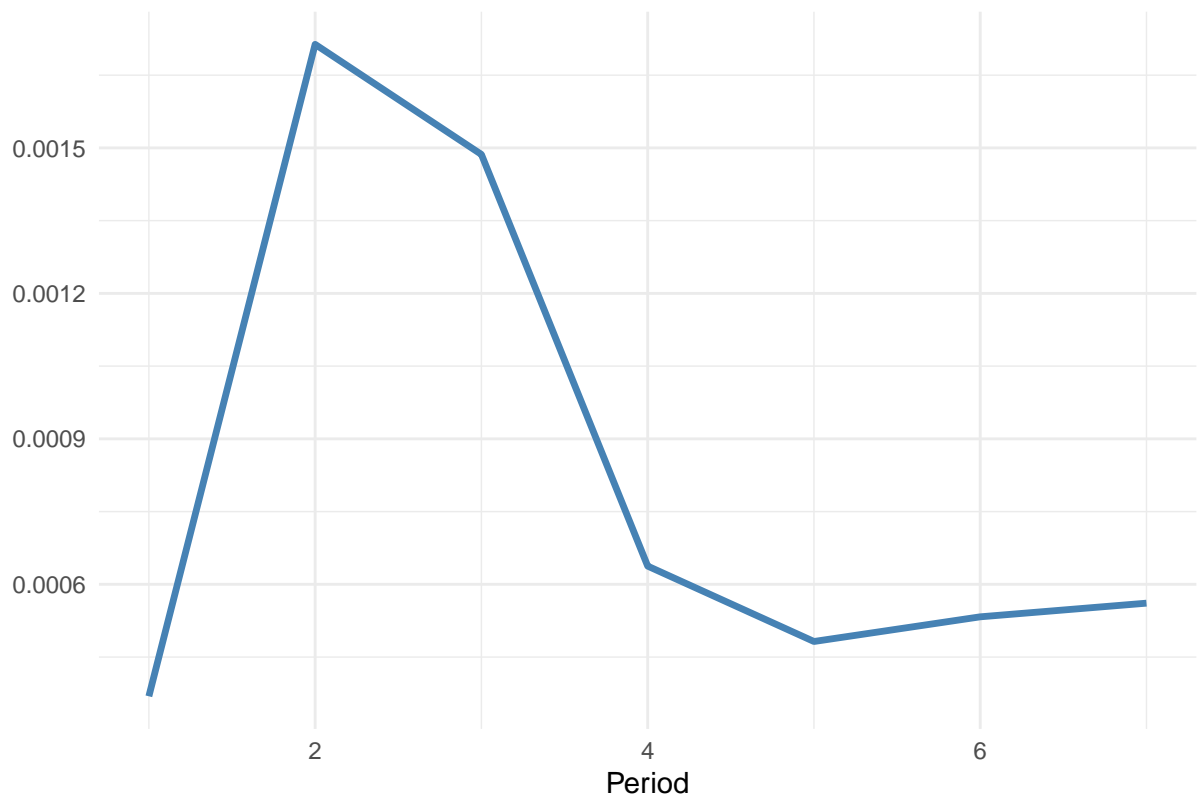


### ARMA-X IRF

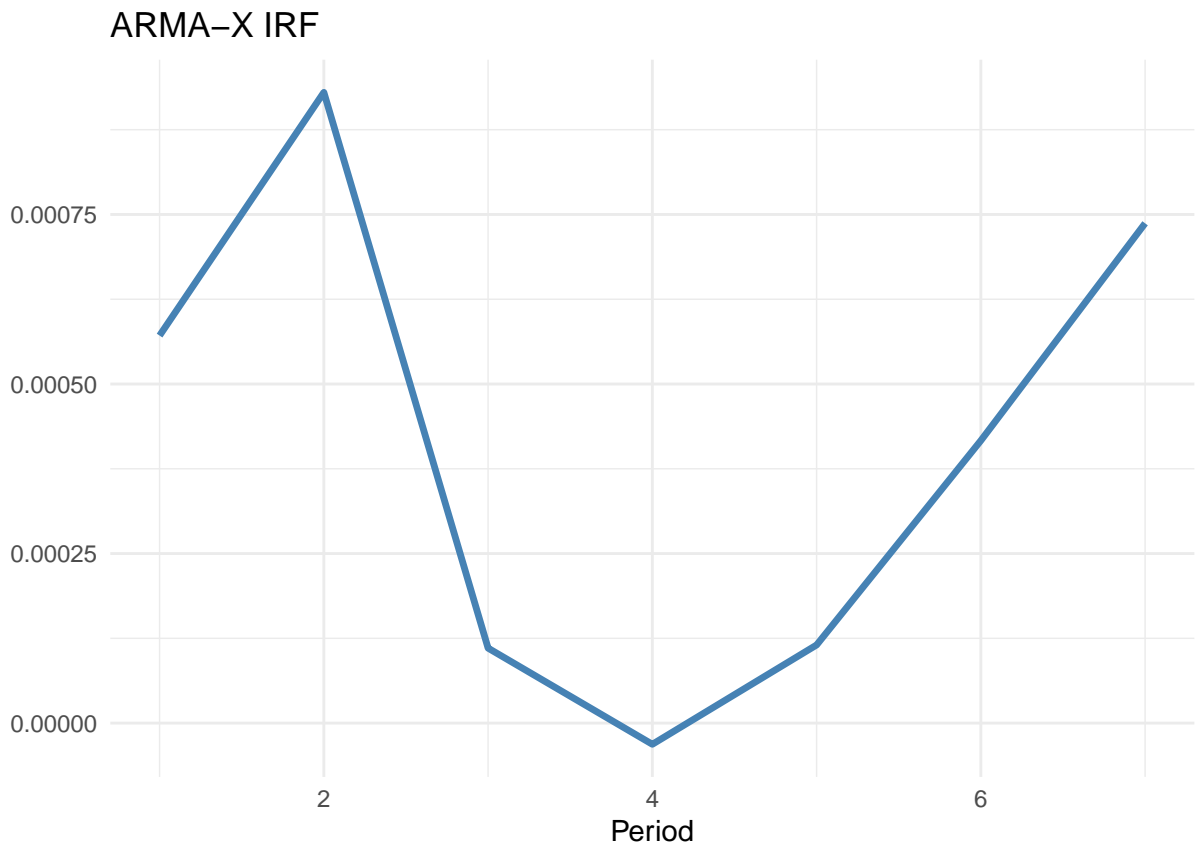


```
irf.plot(res2,nb.periods)
```

ARMA-X IRF



```
irf.plot(res3$model,nb.periods)
```



## Trade Mention as Exogenous

```
#auto.armax selects the lowest AIC value given r (exogenous variable lags)
res1 = auto.armax(data$SPY_vol, xreg=data$trade, nb.lags=2,
                  latex=T, max.p = 6, max.q = 6, max.d=0)
```

```
#armax enables a custom armax specification with p,q,r
res2 = armax(data$SPY_vol, xreg=data$trade, nb.lags=2,
             p=5, q=0, d=0, latex=T)
```

```
#auto.armax.r selects the lowest AIC checking all 3 p,q,r values
res3 = auto.armax.r(data$SPY_vol, x=data$trade,
                   max_p = 6, max_q = 6, max_r = 6, criterion = "AIC", latex=T)
```

```
#we want to plot the IRFs of these models
nb.periods = 7

irf.plot(res1, nb.periods)
```

	Model 1
ar1	0.9828*** (0.0017)
ma1	−0.6774*** (0.0073)
ma2	−0.2149*** (0.0087)
ma3	−0.0126 (0.0080)
ma4	0.0352*** (0.0071)
intercept	0.0222*** (0.0041)
trade_lag_0	0.0019 (0.0019)
trade_lag_1	0.0042* (0.0019)
trade_lag_2	0.0071*** (0.0019)
AIC	−45677.4427
AICc	−45677.4317
BIC	−45598.4233
Log Likelihood	22848.7213
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

Table 10: ARMAX Model Results

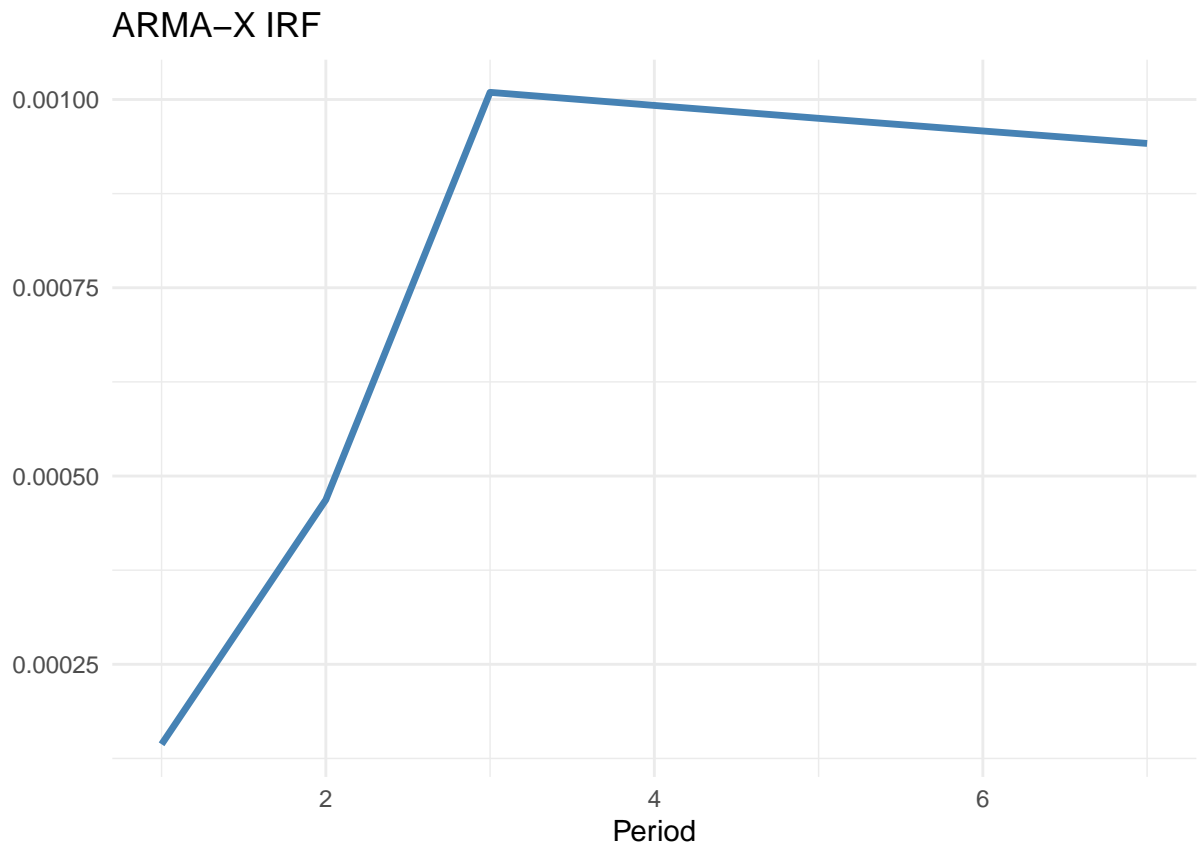
	Model 1
ar1	0.3598*** (0.0071)
ar2	0.0395*** (0.0075)
ar3	0.0974*** (0.0074)
ar4	0.1024*** (0.0075)
ar5	0.0827*** (0.0071)
intercept	0.0221*** (0.0018)
trade_lag_0	0.0027 (0.0019)
trade_lag_1	0.0045* (0.0020)
trade_lag_2	0.0075*** (0.0019)
AIC	−44647.3628
AICc	−44647.3518
BIC	−44568.3435
Log Likelihood	22333.6814
Num. obs.	19969
*** $p < 0.001$ ; ** $p < 0.01$ ; * $p < 0.05$	

Table 11: ARMAX Model Results

	Model 1
ar1	−0.6480*** (0.0101)
ar2	0.0399** (0.0127)
ar3	0.0285*** (0.0085)
ar4	0.1394*** (0.0119)
ar5	0.6351*** (0.0111)
ar6	0.6831*** (0.0076)
ma1	0.9475*** (0.0075)
ma2	0.2704*** (0.0130)
ma3	0.1722*** (0.0112)
ma4	0.0623*** (0.0097)
ma5	−0.6173*** (0.0105)
ma6	−0.8010*** (0.0056)
intercept	0.0225*** (0.0045)
trade_lag_0	0.0006 (0.0017)
trade_lag_1	0.0024 (0.0018)
trade_lag_2	0.0028 (0.0018)
trade_lag_3	−0.0012 (0.0017)
AIC	−47030.2624
AICc	−47030.2281
BIC	−46888.0285
Log Likelihood	23533.1312
Num. obs.	19968

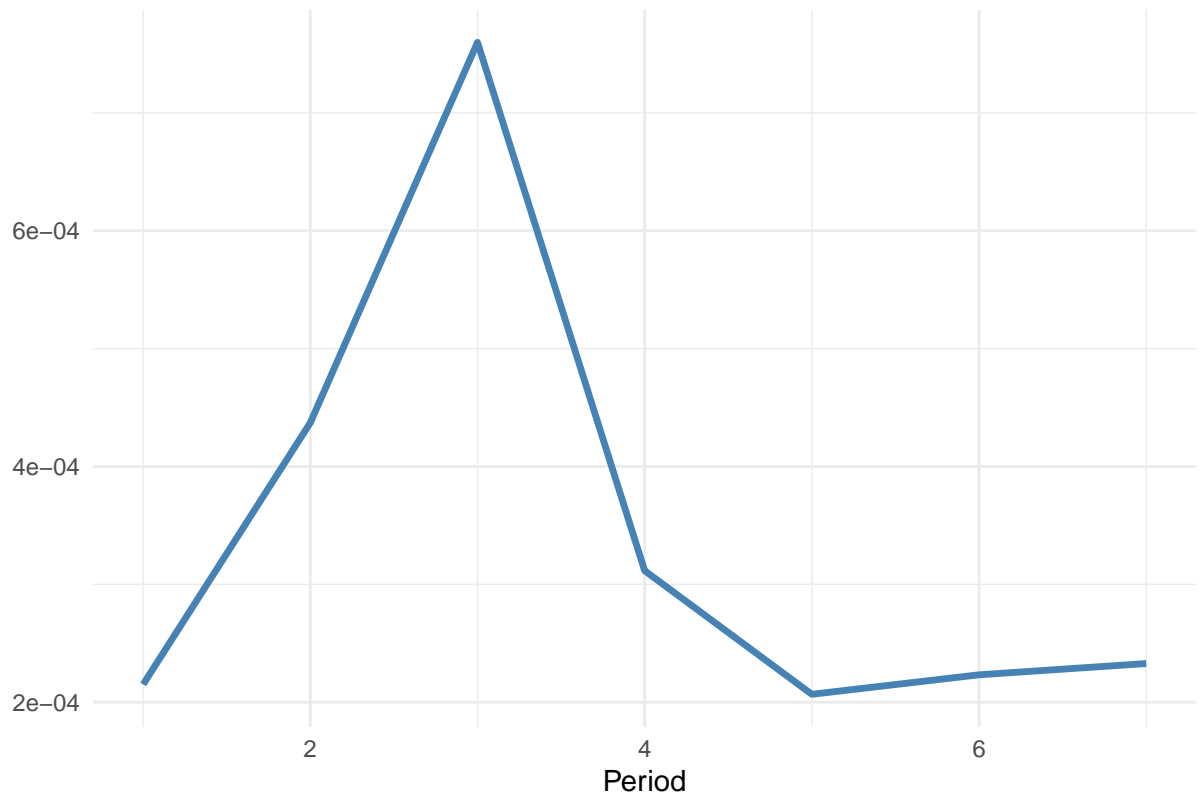
\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

Table 12: ARMAX selected by AIC



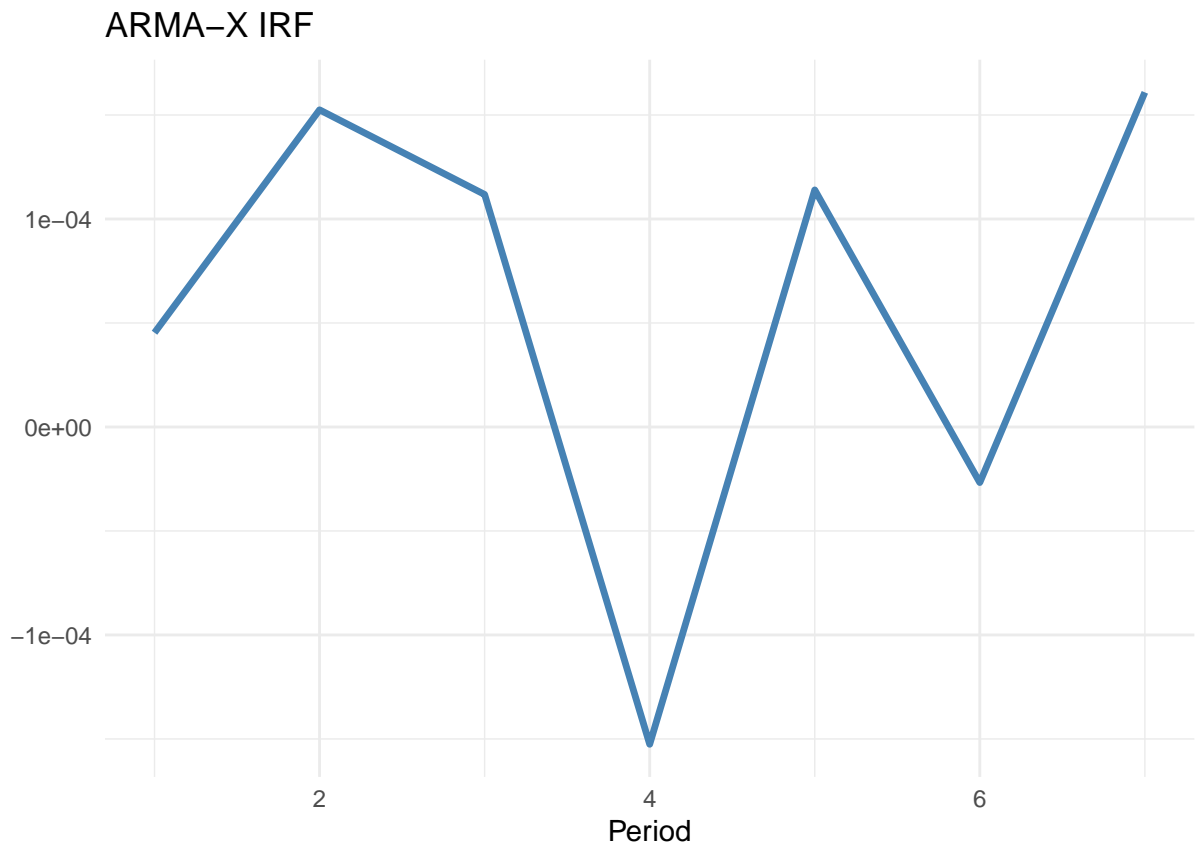
```
irf.plot(res2,nb.periods)
```

### ARMA-X IRF



```
irf.plot(res3$model,nb.periods)
```





## China Mention as Exogenous

```
#auto.armax selects the lowest AIC value given r (exogenous variable lags)
res1 = auto.armax(data$SPY_vol, xreg=data$china, nb.lags=2,
                  latex=T, max.p = 6, max.q = 6, max.d=0)
```

```
#armax enables a custom armax specification with p,q,r
res2 = armax(data$SPY_vol, xreg=data$china, nb.lags=2,
             p=5, q=0, d=0, latex=T)
```

```
#auto.armax.r selects the lowest AIC checking all 3 p,q,r values
res3 = auto.armax.r(data$SPY_vol, x=data$china,
                   max_p = 6, max_q = 6, max_r = 6, criterion = "AIC", latex=T)
```

```
#we want to plot the IRFs of these models
nb.periods = 7

irf.plot(res1, nb.periods)
```

	Model 1
ar1	0.9828*** (0.0017)
ma1	−0.6793*** (0.0073)
ma2	−0.2139*** (0.0087)
ma3	−0.0126 (0.0080)
ma4	0.0355*** (0.0071)
intercept	0.0215*** (0.0041)
china_lag_0	0.0047*** (0.0012)
china_lag_1	0.0084*** (0.0012)
china_lag_2	0.0054*** (0.0012)
AIC	−45721.8164
AICc	−45721.8054
BIC	−45642.7971
Log Likelihood	22870.9082
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

Table 13: ARMAX Model Results

	Model 1
ar1	0.3584*** (0.0071)
ar2	0.0405*** (0.0075)
ar3	0.0972*** (0.0074)
ar4	0.1022*** (0.0075)
ar5	0.0825*** (0.0071)
intercept	0.0215*** (0.0018)
china_lag_0	0.0047*** (0.0012)
china_lag_1	0.0079*** (0.0013)
china_lag_2	0.0051*** (0.0012)
AIC	−44680.9095
AICc	−44680.8985
BIC	−44601.8902
Log Likelihood	22350.4548
Num. obs.	19969
*** $p < 0.001$ ; ** $p < 0.01$ ; * $p < 0.05$	

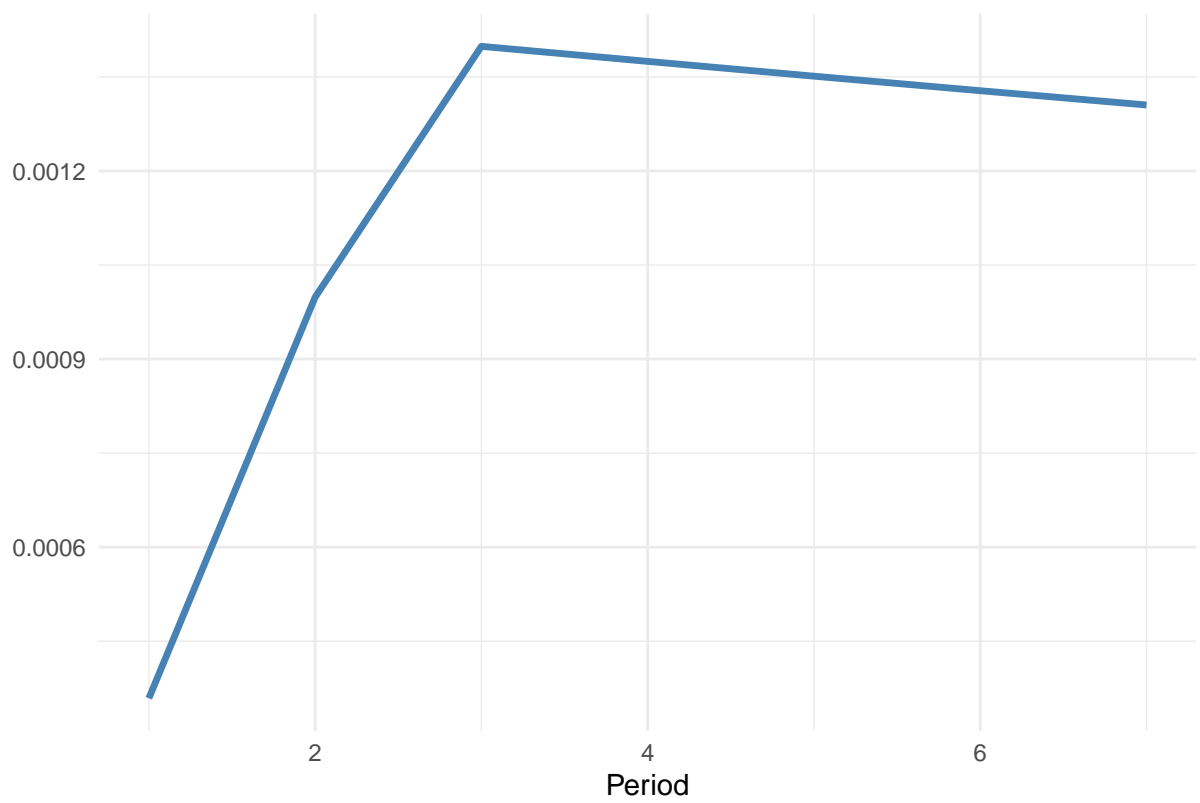
Table 14: ARMAX Model Results

	Model 1
ar1	−0.6497*** (0.0103)
ar2	0.0325* (0.0132)
ar3	0.0203* (0.0086)
ar4	0.1350*** (0.0120)
ar5	0.6374*** (0.0110)
ar6	0.6859*** (0.0076)
ma1	0.9502*** (0.0078)
ma2	0.2775*** (0.0136)
ma3	0.1797*** (0.0118)
ma4	0.0682*** (0.0101)
ma5	−0.6136*** (0.0106)
ma6	−0.7998*** (0.0056)
intercept	0.0217*** (0.0040)
china_lag_0	0.0039*** (0.0011)
china_lag_1	0.0060*** (0.0011)
china_lag_2	0.0049*** (0.0011)
AIC	−47085.9168
AICc	−47085.8861
BIC	−46951.5839
Log Likelihood	23559.9584
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

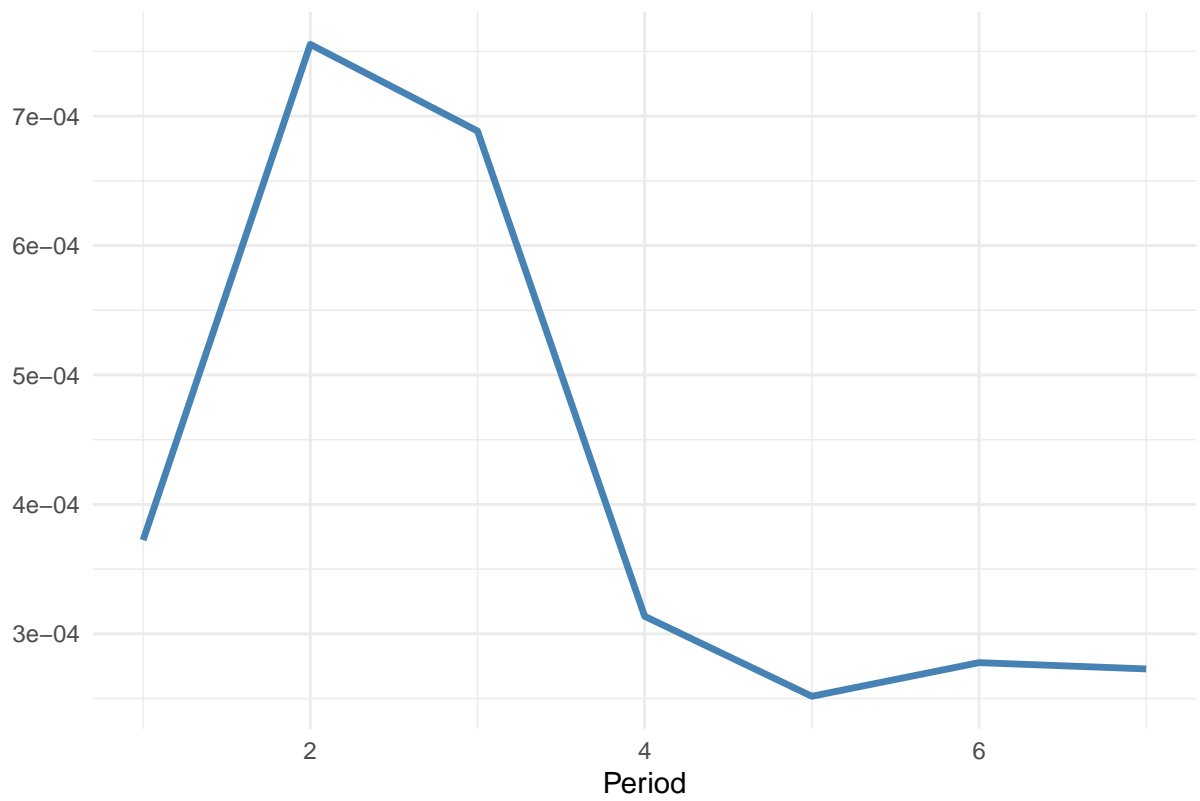
Table 15: ARMAX selected by AIC

ARMA-X IRF

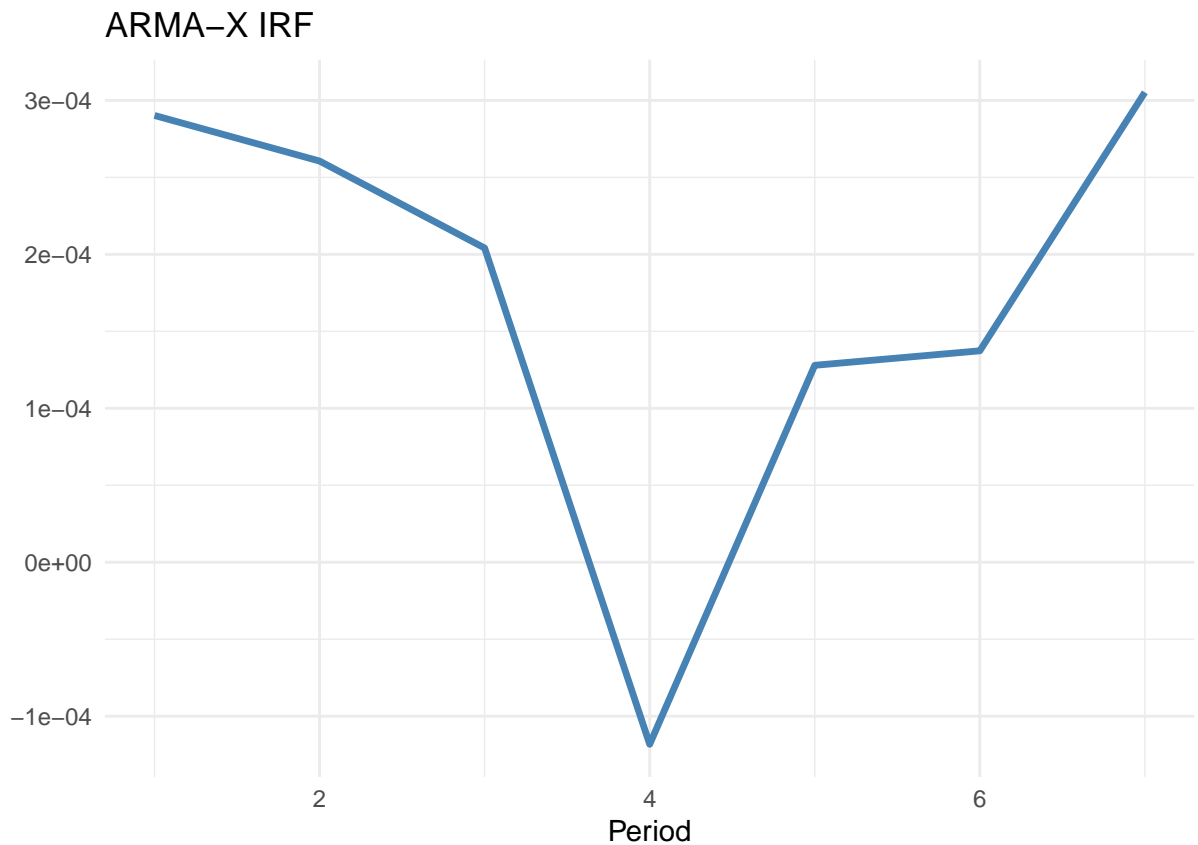


```
irf.plot(res2,nb.periods)
```

ARMA-X IRF



```
irf.plot(res3$model,nb.periods)
```



## Positive Vibe as Exogenous

```
#auto.armax selects the lowest AIC value given r (exogenous variable lags)
res1 = auto.armax(data$SPY_vol, xreg=data$prop_positive, nb.lags=2,
                  latex=T, max.p = 6, max.q = 6, max.d=0)
```

```
#armax enables a custom armax specification with p,q,r
res2 = armax(data$SPY_vol, xreg=data$prop_positive, nb.lags=2,
             p=5, q=0, d=0, latex=T)
```

```
#auto.armax.r selects the lowest AIC checking all 3 p,q,r values
res3 = auto.armax.r(data$SPY_vol, x=data$prop_positive,
                   max_p = 6, max_q = 6, max_r = 6, criterion = "AIC", latex=T)
```

```
#we want to plot the IRFs of these models
nb.periods = 7

irf.plot(res1, nb.periods)
```

	Model 1
ar1	0.9828*** (0.0017)
ma1	−0.6777*** (0.0073)
ma2	−0.2146*** (0.0087)
ma3	−0.0117 (0.0080)
ma4	0.0345*** (0.0071)
intercept	0.0195*** (0.0042)
prop_positive_lag_0	0.0068*** (0.0017)
prop_positive_lag_1	0.0027 (0.0017)
prop_positive_lag_2	0.0045** (0.0017)
AIC	−45686.3386
AICc	−45686.3275
BIC	−45607.3192
Log Likelihood	22853.1693
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

Table 16: ARMAX Model Results



	Model 1
ar1	0.3594*** (0.0071)
ar2	0.0398*** (0.0075)
ar3	0.0987*** (0.0074)
ar4	0.1032*** (0.0075)
ar5	0.0820*** (0.0071)
intercept	0.0194*** (0.0019)
prop_positive_lag_0	0.0073*** (0.0017)
prop_positive_lag_1	0.0030 (0.0017)
prop_positive_lag_2	0.0045** (0.0017)
AIC	−44656.4789
AICc	−44656.4679
BIC	−44577.4595
Log Likelihood	22338.2394
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

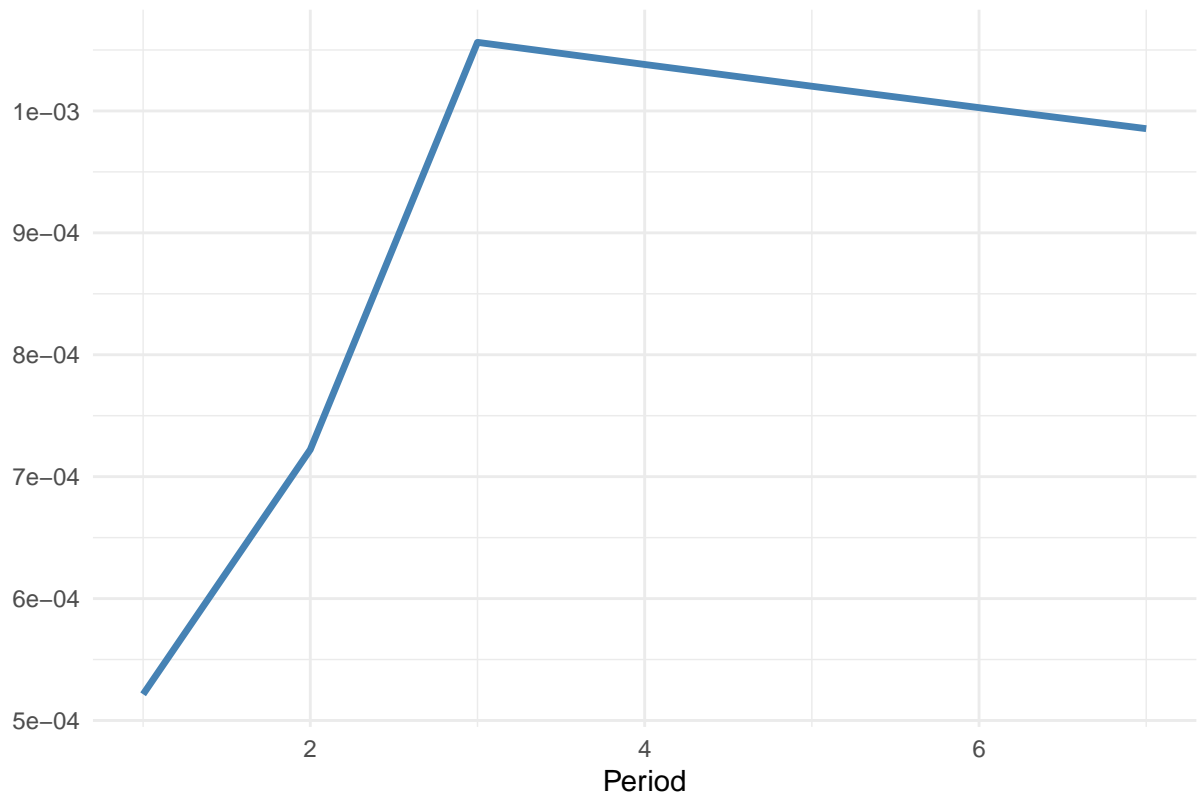
Table 17: ARMAX Model Results

	Model 1
ar1	−0.6498*** (0.0102)
ar2	0.0305* (0.0129)
ar3	0.0197* (0.0085)
ar4	0.1384*** (0.0118)
ar5	0.6398*** (0.0109)
ar6	0.6852*** (0.0076)
ma1	0.9498*** (0.0077)
ma2	0.2786*** (0.0132)
ma3	0.1814*** (0.0115)
ma4	0.0670*** (0.0099)
ma5	−0.6165*** (0.0104)
ma6	−0.8008*** (0.0055)
intercept	0.0204*** (0.0042)
prop_positive_lag_0	0.0041** (0.0015)
prop_positive_lag_1	0.0034* (0.0016)
prop_positive_lag_2	0.0025 (0.0015)
AIC	−47042.9836
AICc	−47042.9529
BIC	−46908.6507
Log Likelihood	23538.4918
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

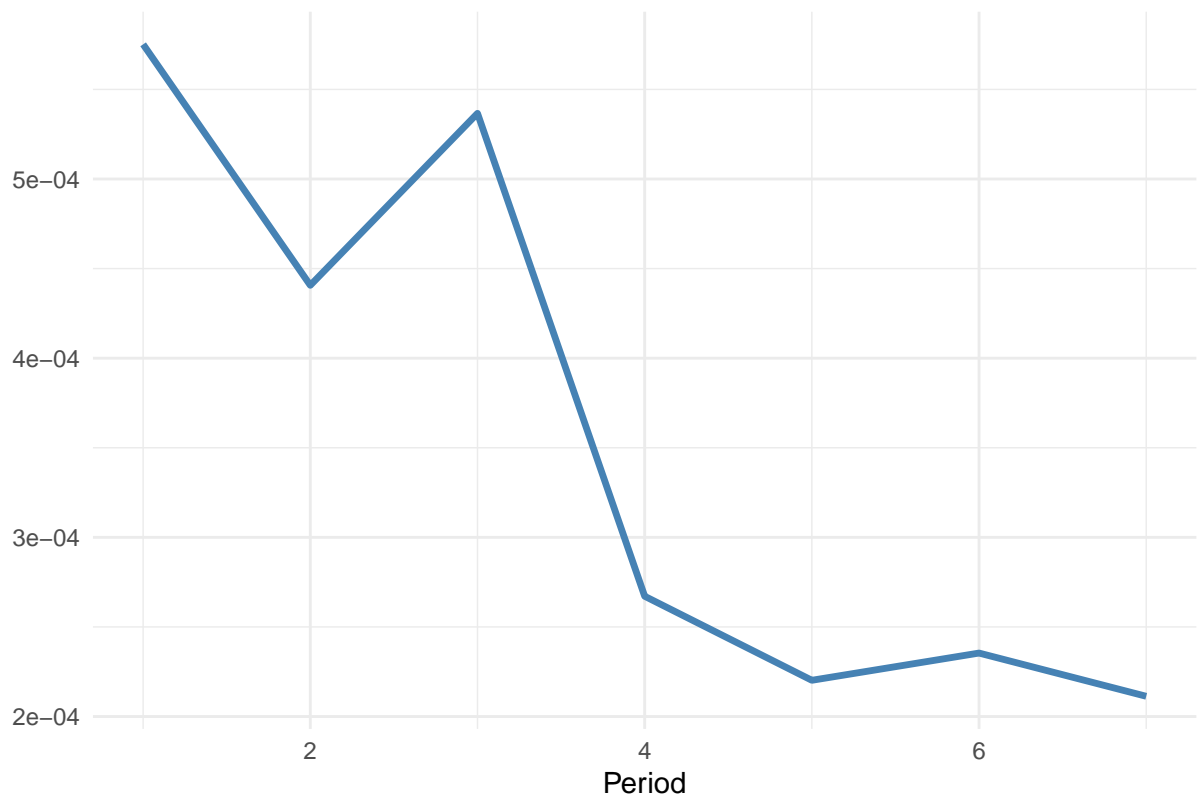
Table 18: ARMAX selected by AIC

ARMA-X IRF

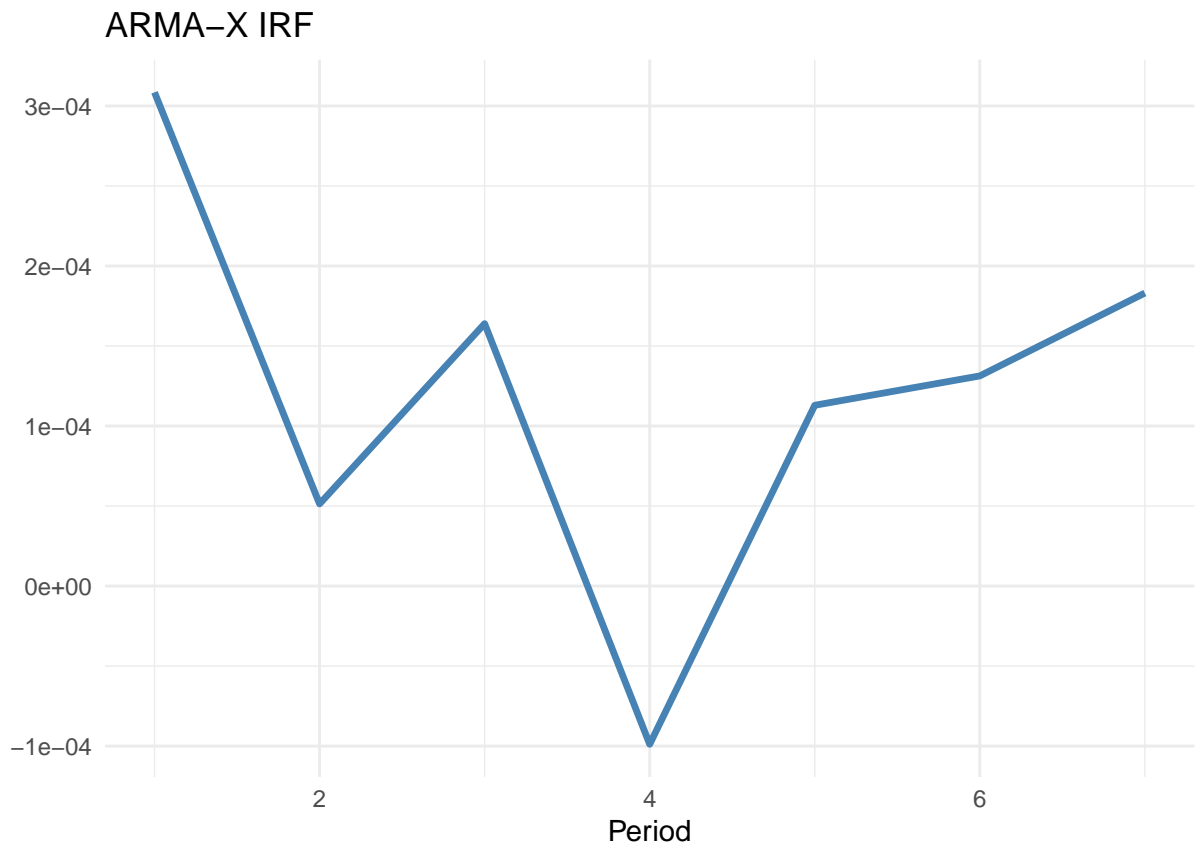


```
irf.plot(res2,nb.periods)
```

ARMA-X IRF



```
irf.plot(res3$model,nb.periods)
```



## Negative Vibe as Exogenous

```
#auto.armax selects the lowest AIC value given r (exogenous variable lags)
res1 = auto.armax(data$SPY_vol, xreg=data$prop_negative, nb.lags=2,
                  latex=T, max.p = 6, max.q = 6, max.d=0)
```

```
#arimax enables a custom arimax specification with p,q,r
res2 = arimax(data$SPY_vol, xreg=data$prop_negative, nb.lags=2,
              p=5, q=0, d=0, latex=T)
```

```
#auto.armax.r selects the lowest AIC checking all 3 p,q,r values
res3 = auto.armax.r(data$SPY_vol, x=data$prop_negative,
                    max_p = 6, max_q = 6, max_r = 6, criterion = "AIC", latex=T)
```

```
#we want to plot the IRFs of these models
nb.periods = 7

irf.plot(res1, nb.periods)
```

	Model 1
ar1	0.9828*** (0.0017)
ma1	−0.6776*** (0.0073)
ma2	−0.2142*** (0.0087)
ma3	−0.0126 (0.0080)
ma4	0.0348*** (0.0071)
intercept	0.0213*** (0.0042)
prop_negative_lag_0	0.0068** (0.0023)
prop_negative_lag_1	0.0030 (0.0023)
prop_negative_lag_2	−0.0007 (0.0023)
AIC	−45673.0125
AICc	−45673.0015
BIC	−45593.9932
Log Likelihood	22846.5063
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

Table 19: ARMAX Model Results

	Model 1
ar1	0.3594*** (0.0071)
ar2	0.0403*** (0.0075)
ar3	0.0978*** (0.0074)
ar4	0.1033*** (0.0075)
ar5	0.0819*** (0.0071)
intercept	0.0212*** (0.0019)
prop_negative_lag_0	0.0075** (0.0023)
prop_negative_lag_1	0.0036 (0.0023)
prop_negative_lag_2	−0.0012 (0.0023)
AIC	−44643.1033
AICc	−44643.0923
BIC	−44564.0840
Log Likelihood	22331.5517
Num. obs.	19969

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

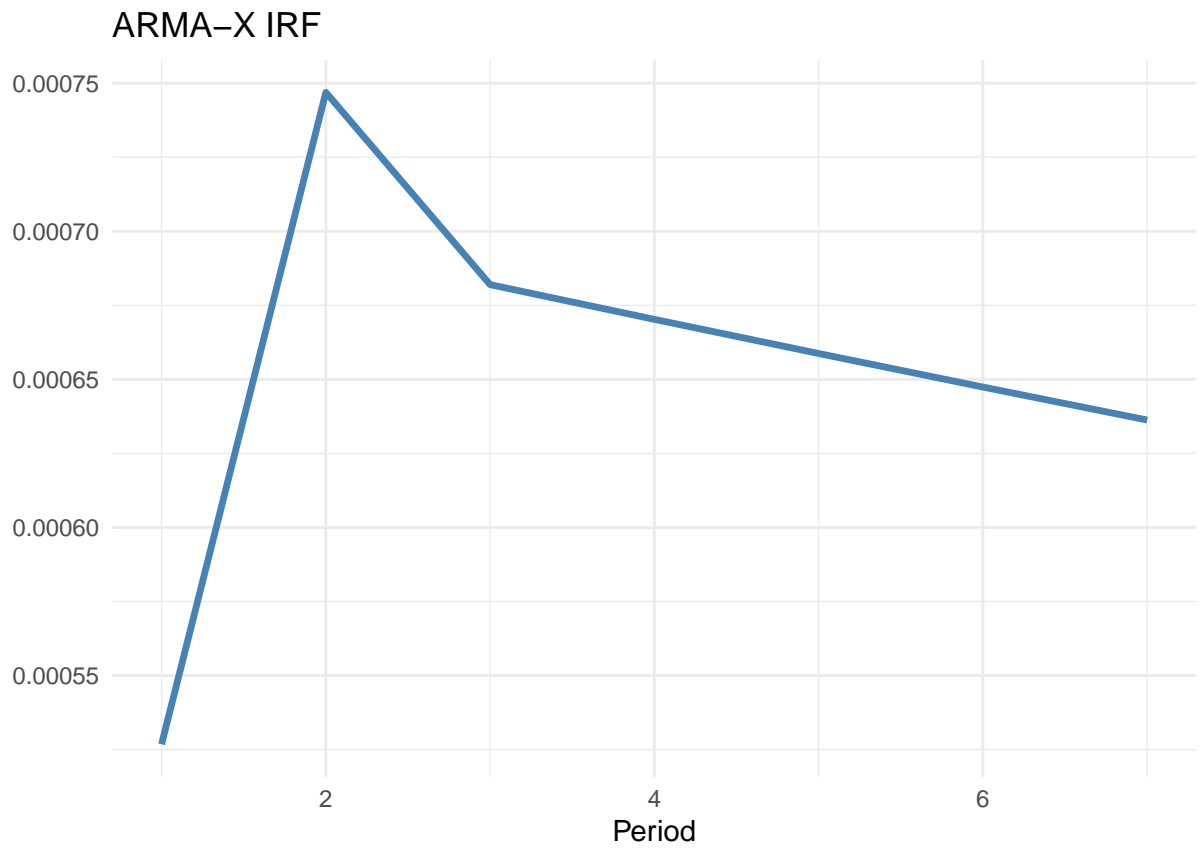
Table 20: ARMAX Model Results

	Model 1
ar1	−0.6518*** (0.0101)
ar2	0.0293* (0.0129)
ar3	0.0196* (0.0084)
ar4	0.1367*** (0.0117)
ar5	0.6404*** (0.0108)
ar6	0.6871*** (0.0075)
ma1	0.9511*** (0.0076)
ma2	0.2807*** (0.0132)
ma3	0.1829*** (0.0115)
ma4	0.0688*** (0.0100)
ma5	−0.6155*** (0.0104)
ma6	−0.8013*** (0.0055)
intercept	0.0213*** (0.0041)
prop_negative_lag_0	0.0061** (0.0020)
prop_negative_lag_1	0.0034 (0.0020)
AIC	−47047.3129
AICc	−47047.2856
BIC	−46920.8811
Log Likelihood	23539.6564
Num. obs.	19970

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

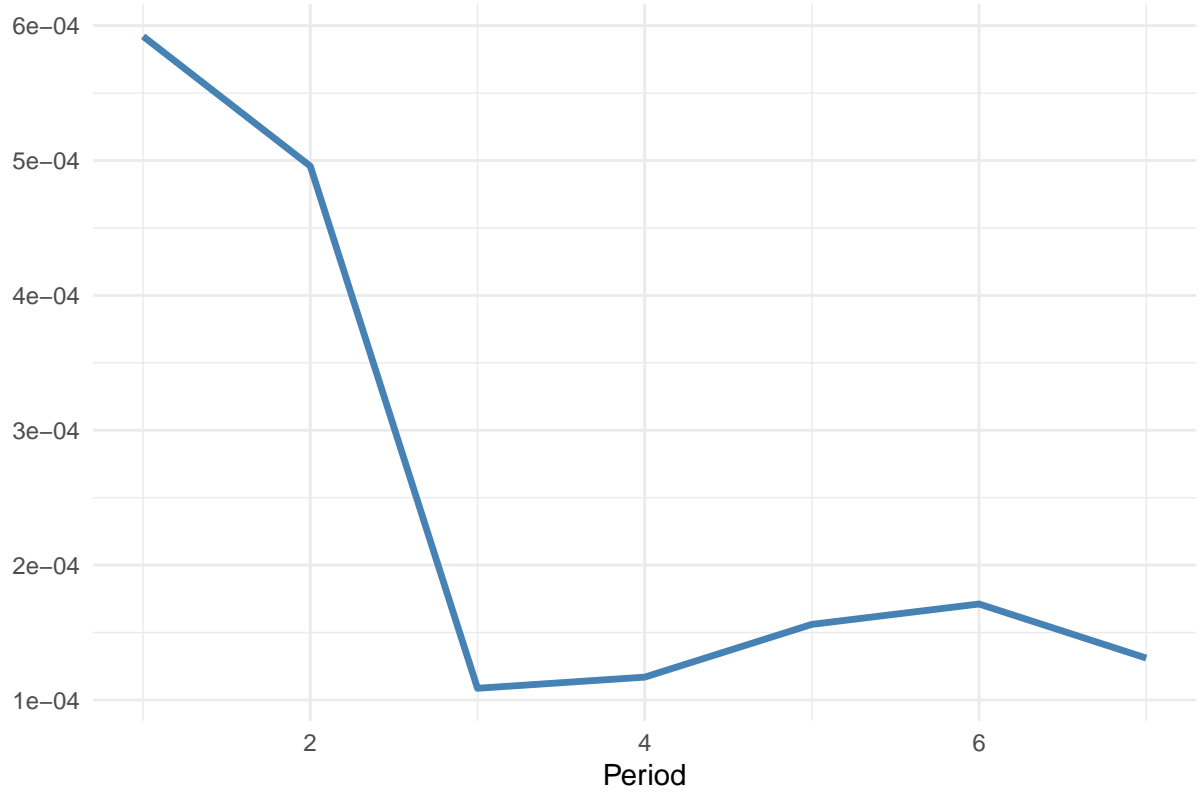
Table 21: ARMAX selected by AIC





```
irf.plot(res2,nb.periods)
```

### ARMA-X IRF



```
irf.plot(res3$model,nb.periods)
```

