ARMA-X Figures

Contents

Full Timeframe (Jan 2024 to May 2025)

```r
#load final dataset
source(here("helperfunctions/full_data.R"))

#backup
backup = data

#select timeframe
data = filter(data,between(timestamp, as.Date('2014-01-01'), as.Date('2025-05-07')))

#for interpretation
mean1 = mean(data$SPY_vol)
```

SPY Models

We choose the specification in the armax_models file. In this file, we will just run said specifications to produce nice tables and graphs to include in our final paper. This is also why there are specification differences in the separate timeframes. We always use the best fit we found earlier.

```r
models <- list()

# ARMA-X(3,3,1) with Tweet Dummy as Exogenous
models[["Model 1"]] <- armax(data$SPY_vol, xreg = data$dummy, latex = F,
                             nb.lags = 1, p = 3, q = 3)

# ARMA-X(3,3,1) with Tweet Count as Exogenous
models[["Model 2"]] <- armax(data$SPY_vol, xreg = data$N, latex = F,
                             nb.lags = 1, p = 3, q = 3)

# ARMA-X(3,2,3) with Tariff Mentions as Exogenous
models[["Model 3"]] <- armax(data$SPY_vol, xreg = data$tariff, latex = F,
                             nb.lags = 3, p = 3, q = 2)

# ARMA-X(3,2,1) with Trade Mentions as Exogenous
models[["Model 4"]] <- armax(data$SPY_vol, xreg = data$trade, latex = F,
                             nb.lags = 1, p = 3, q = 2)

# ARMA-X(3,2,0) with China Mentions as Exogenous
models[["Model 5"]] <- armax(data$SPY_vol, xreg = data$china, latex = F,
                             nb.lags = 0, p = 3, q = 2)
```

SPY Table

```r
names = list( "ar1" = "AR(1)",
              "ar2" = "AR(2)",
              "ar3" = "AR(3)",
              "ma1" = "MA(1)",
              "ma2" = "MA(2)",
```

```r
              "ma3" = "MA(3)",
              "(Intercept)" = "Constant",
              "dummy_lag_0" = "$TweetDummy_{t}$",
              "dummy_lag_1" = "$TweetDummy_{t-1}$",
              "N_lag_0" = "$TweetCount_{t}$",
              "N_lag_1" = "$TweetCount_{t-1}$",
              "tariff_lag_0" = "$Tariff_{t}$",
              "tariff_lag_1" = "$Tariff_{t-1}$",
              "tariff_lag_2" = "$Tariff_{t-2}$",
              "tariff_lag_3" = "$Tariff_{t-3}$",
              "trade_lag_0" = "$Trade_{t}$",
              "trade_lag_1" = "$Trade_{t-1}$",
              "china_lag_0" = "$China_{t}$")

texreg(models,
          custom.model.names = names(models),
          custom.coef.map = names,
          caption = "ARMAX Models of Average Hourly Volatility",
          caption.above = TRUE,
          label = "tab:armax",
          digits = 4)
```

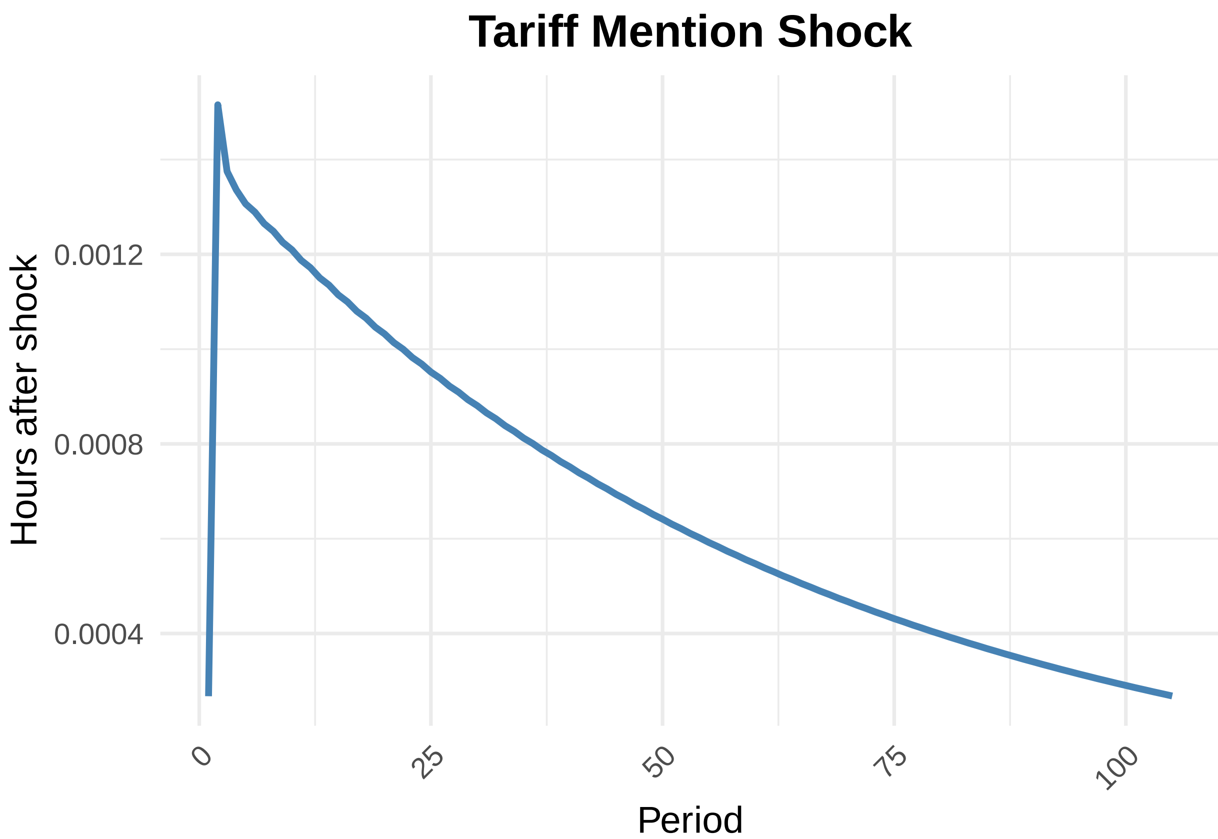
SPY IRFs

```r
#we want to plot the IRFs of these models
nb.periods = 7 * 15

#irf.plot(models[["Model 1"]],nb.periods,title="Tweet Dummy Shock")
#irf.plot(models[["Model 2"]],nb.periods,title="Tweet Count Shock")
irf.plot(models[["Model 3"]],nb.periods,title="Tariff Mention Shock")
```

Table 1: ARMAX Models of Average Hourly Volatility

	Model 1	Model 2	Model 3	Model 4	Model 5
AR(1)	0.0300	0.0278	0.2200^{***}	2.1903^{***}	0.2209^{***}
	(0.0510)	(0.0510)	(0.0084)	(0.0096)	(0.0084)
AR(2)	0.7229^{***}	0.7210^{***}	0.9388^{***}	-1.4727^{***}	0.9382^{***}
	(0.0397)	(0.0399)	(0.0037)	(0.0173)	(0.0037)
AR(3)	0.2110^{***}	0.2148^{***}	-0.1837^{***}	0.2784^{***}	-0.1837^{***}
	(0.0287)	(0.0284)	(0.0079)	(0.0082)	(0.0079)
MA(1)	0.2751^{***}	0.2779^{***}	0.0870^{***}	-1.8955^{***}	0.0878^{***}
	(0.0496)	(0.0496)	(0.0042)	(0.0062)	(0.0042)
MA(2)	-0.6445^{***}	-0.6430^{***}	-0.8960^{***}	0.9165^{***}	-0.8950^{***}
	(0.0284)	(0.0285)	(0.0042)	(0.0063)	(0.0042)
MA(3)	-0.3527^{***}	-0.3563^{***}			
	(0.0256)	(0.0253)			
$TweetDummy_t$	0.0014^{***}				
	(0.0002)				
$TweetDummy_{t-1}$	0.0008^{***}				
	(0.0002)				
$TweetCount_t$		0.0004^{***}			
		(0.0001)			
$TweetCount_{t-1}$		0.0002^{**}			
		(0.0001)			
$Tariff_t$			0.0035^{*}		
			(0.0014)		
$Tariff_{t-1}$			0.0191^{***}		
			(0.0015)		
$Tariff_{t-2}$			0.0103^{***}		
			(0.0015)		
$Tariff_{t-3}$			-0.0045^{**}		
			(0.0014)		
$Trade_t$				0.0032	
				(0.0018)	
$Trade_{t-1}$				0.0016	
				(0.0018)	
$China_t$					0.0026^{*}
					(0.0012)
AIC	−45761.2161	−45737.6695	−46020.9547	−45816.1540	−45840.5349
AICc	−45761.2051	−45737.6585	−46020.9415	−45816.1449	−45840.5277
BIC	−45682.1963	−45658.6497	−45934.0340	−45745.0361	−45777.3186
Log Likelihood	22890.6081	22878.8348	23021.4774	22917.0770	22928.2675
Num. obs.	19970	19970	19968	19970	19971

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$




```
#irf.plot(models[["Model 4"]],nb.periods,title="Trade Mention Shock")
#irf.plot(models[["Model 5"]],nb.periods,title="China Mention Shock")
```

SPY Residuals

```
res1 = checkresiduals(models[["Model 1"]], plot = FALSE)
res2 = checkresiduals(models[["Model 2"]], plot = FALSE)
res3 = checkresiduals(models[["Model 3"]], plot = FALSE)
res4 = checkresiduals(models[["Model 4"]], plot = FALSE)
res5 = checkresiduals(models[["Model 5"]], plot = FALSE)
```

```
resnames = c("Twitter Dummy", "Twitter Count", "Tariff", "Trade", "China")

#extract p-values directly from checkresiduals results
pvals <- data.frame(Model = resnames,
                    `Ljung-Box p-value` = c(
                      res1$p.value,
                      res2$p.value,
                      res3$p.value,
                      res4$p.value,
                      res5$p.value))

#table
knitr::kable(pvals, digits = 100, caption = "Full Timeframe Ljung-Box Test p-values")
```

Table 2: Full Timeframe Ljung-Box Test p-values

Model	Ljung.Box.p.value
Twitter Dummy	0
Twitter Count	0
Tariff	0
Trade	0
China	0

First Term

```
#load final dataset
data = backup

#first term
data = filter(data,between(timestamp, as.Date('2017-01-20'), as.Date('2021-01-20')))

#for interpretation
mean2 = mean(data$SPY_vol)
```

SPY Models

```
models <- list()

# ARMA-X(3,3,0) with Tariff Mentions as Exogenous
models[["First Term (1)"]] <- armax(data$SPY_vol, xreg = data$tariff, latex = F,
                              nb.lags = 0, p = 3, q = 3)

# ARMA-X(3,3,0) with Trade Mentions as Exogenous
models[["First Term (2)"]] <- armax(data$SPY_vol, xreg = data$trade, latex = F,
                              nb.lags = 0, p = 3, q = 3)

# ARMA-X(3,3,0) with Trade Mentions as Exogenous
models[["First Term (3)"]] <- armax(data$SPY_vol, xreg = data$china, latex = F,
                              nb.lags = 0, p = 3, q = 3)
```

SPY Residuals

```
res6 = checkresiduals(models[["First Term (1)"]], plot = FALSE)
res7 = checkresiduals(models[["First Term (2)"]], plot = FALSE)
res8 = checkresiduals(models[["First Term (3)"]], plot = FALSE)

pvals_new1 <- data.frame(
  Model = c("First Term Tariffs", "First Term Trade", "First Term China"),
```

```r
  `Ljung-Box p-value` = c(
    res6$p.value,
    res7$p.value,
    res8$p.value))
```

Second Term

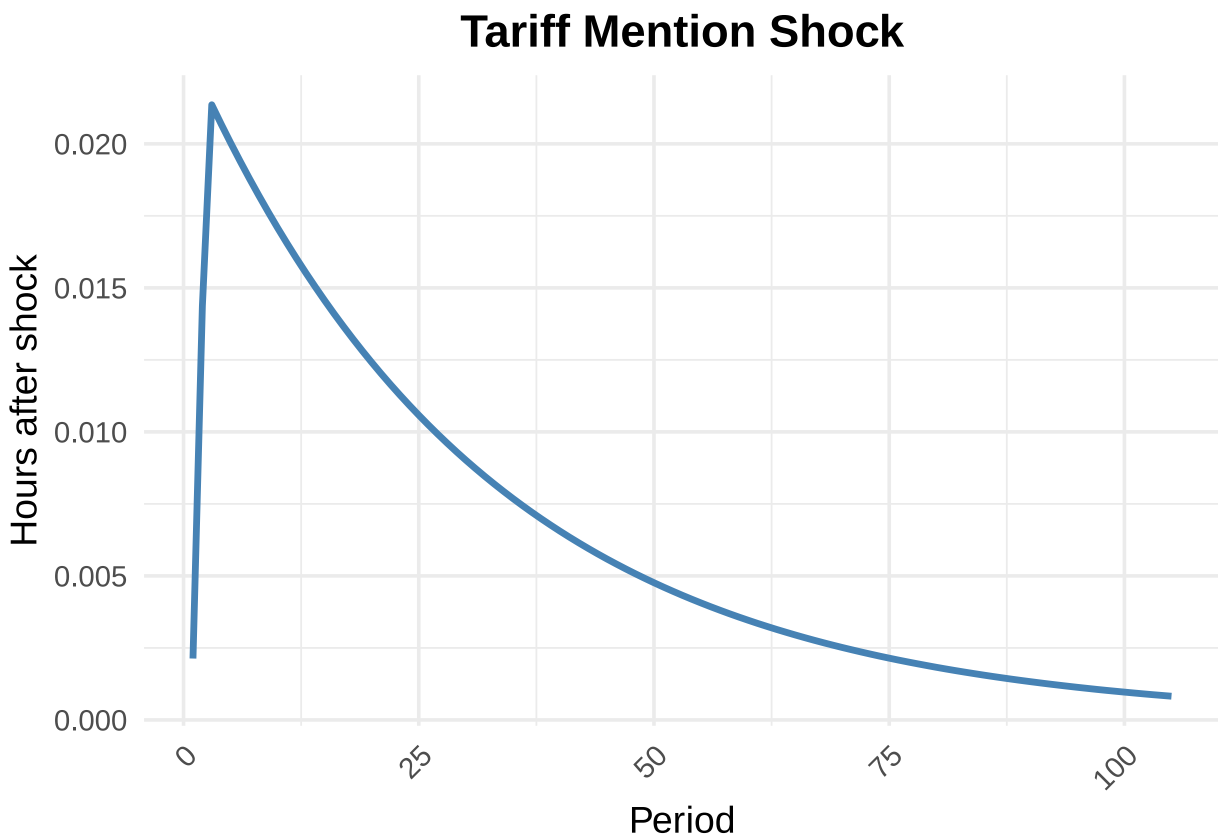
```r
#load final dataset
data = backup

#second term
data = filter(data,between(timestamp, as.Date('2025-01-20'), as.Date('2025-05-07')))

#for interpretation
mean3 = mean(data$SPY_vol)
```

SPY Models

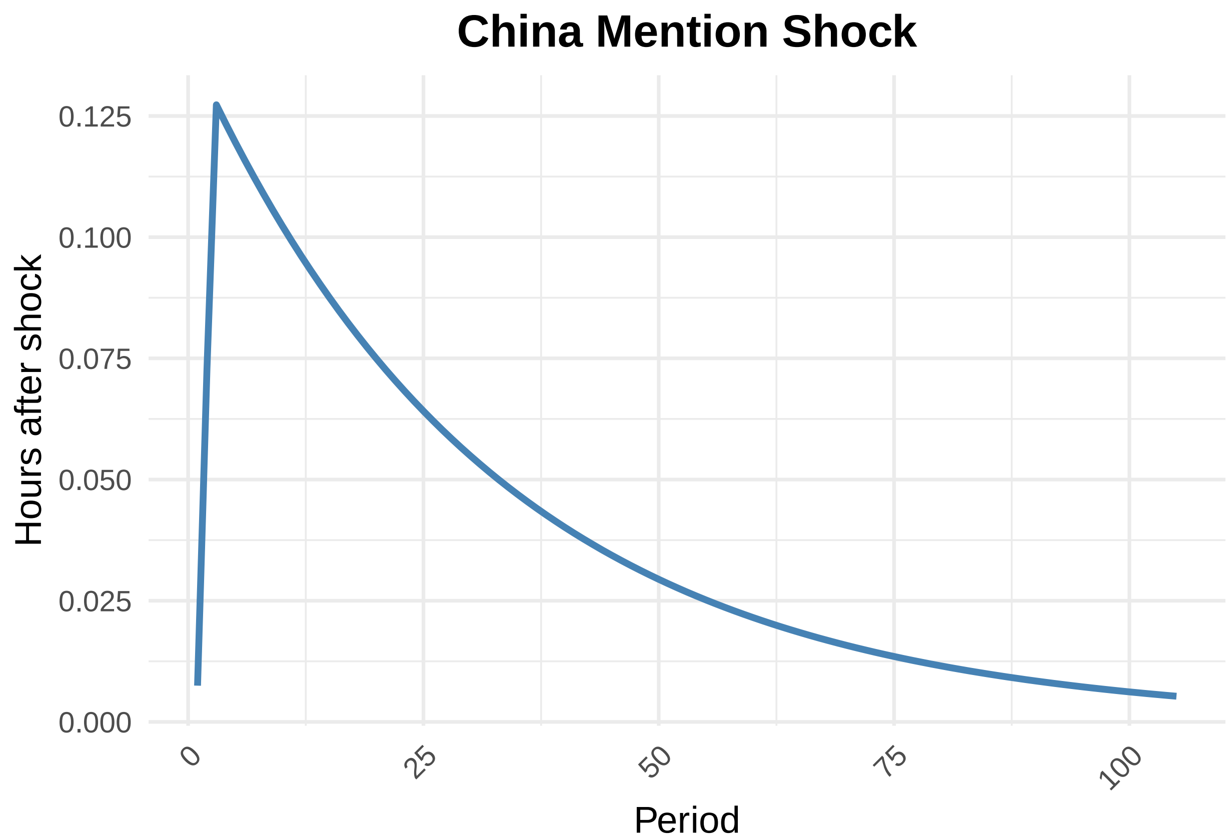
```r
# ARMA-X(3,2,3) with Tariff Mentions as Exogenous
models[["Second Term (1)"]] <- armax(data$SPY_vol, xreg = data$tariff, latex = F,
                             nb.lags = 2, p = 1, q = 2)

# ARMA-X(3,2,1) with Trade Mentions as Exogenous
models[["Second Term (2)"]] <- armax(data$SPY_vol, xreg = data$trade, latex = F,
                             nb.lags = 0, p = 1, q = 2)

# ARMA-X(3,2,0) with China Mentions as Exogenous
models[["Second Term (3)"]] <- armax(data$SPY_vol, xreg = data$china, latex = F,
                             nb.lags = 2, p = 1, q = 2)
```

SPY IRFs

```r
#we want to plot the IRFs of these models
nb.periods = 7 * 15

irf.plot(models[["Second Term (1)"]],nb.periods,title="Tariff Mention Shock")
```




```
irf.plot(models[["Second Term (3)"]],nb.periods,title="China Mention Shock")
```



SPY Residuals

```
res9 = checkresiduals(models[["Second Term (1)"]], plot = FALSE)
res10 = checkresiduals(models[["Second Term (2)"]], plot = FALSE)
res11 = checkresiduals(models[["Second Term (3)"]], plot = FALSE)

pvals_new2 <- data.frame(
  Model = c("Second Term Tariffs", "Second Term Trade", "Second Term China"),
  `Ljung-Box p-value` = c(
    res9$p.value,
    res10$p.value,
    res11$p.value))

#combine with other term
pvals_combined <- rbind(pvals_new1, pvals_new2)
```

SPY Table (both terms)

```
xnames = list("ar1" = "AR(1)",
              "ar2" = "AR(2)",
              "ar3" = "AR(3)",
              "ma1" = "MA(1)",
```

```r
              "ma2" = "MA(2)",
              "ma3" = "MA(3)",
              "(Intercept)" = "Constant",
              "tariff_lag_0" = "$Tariff_{t}$",
              "tariff_lag_1" = "$Tariff_{t-1}$",
              "tariff_lag_2" = "$Tariff_{t-2}$",
              "trade_lag_0" = "$Trade_{t}$",
              "china_lag_0" = "$China_{t}$",
              "china_lag_1" = "$China_{t-1}$",
              "china_lag_2" = "$China_{t-2}$")

texreg(models,
       custom.model.names = names(models),
       custom.coef.map = xnames,
       caption = "Split-Term ARMAX Models of Average Hourly Volatility",
       caption.above = TRUE,
       label = "tab:armax_term",
       digits = 4)
```

SPY Residuals Table (both terms)

```r
knitr::kable(pvals_combined, digits = 100, caption = "Separate Terms Ljung-Box Test p-values")
```

Table 4: Separate Terms Ljung-Box Test p-values

Model	Ljung.Box.p.value
First Term Tariffs	0.0000000
First Term Trade	0.0000000
First Term China	0.0000000
Second Term Tariffs	0.8489828
Second Term Trade	0.8322070
Second Term China	0.5122385

Descriptive Stats

```r
means <- data.frame(
  Model = c("Full Time Mean", "First Term Mean", "Second Term Mean"),
  `SPY Volatility Mean` = c(
    mean1,
    mean2,
    mean3))

knitr::kable(means, digits = 6, caption = "Summary Statistics of SPY Volatility")
```

Table 3: Split-Term ARMAX Models of Average Hourly Volatility

	First Term (1)	First Term (2)	First Term (3)	Second Term (1)	Second Term (2)	Second Term (
AR(1)	0.2953^{***}	0.2943^{***}	0.2927^{***}	0.9686^{***}	0.9683^{***}	0.9693^{***}
	(0.0225)	(0.0224)	(0.0224)	(0.0163)	(0.0163)	(0.0161)
AR(2)	0.1434^{***}	0.1439^{***}	0.1438^{***}			
	(0.0220)	(0.0220)	(0.0219)			
AR(3)	0.5456^{***}	0.5462^{***}	0.5480^{***}			
	(0.0223)	(0.0222)	(0.0222)			
MA(1)	0.1854^{***}	0.1863^{***}	0.1866^{***}	-0.6965^{***}	-0.6905^{***}	-0.7207^{***}
	(0.0180)	(0.0179)	(0.0179)	(0.0469)	(0.0469)	(0.0467)
MA(2)	-0.1707^{***}	-0.1706^{***}	-0.1695^{***}	-0.1732^{***}	-0.1755^{***}	-0.1609^{***}
	(0.0169)	(0.0169)	(0.0168)	(0.0437)	(0.0438)	(0.0434)
MA(3)	-0.6557^{***}	-0.6564^{***}	-0.6575^{***}			
	(0.0162)	(0.0161)	(0.0161)			
$Tariff_t$	0.0011			0.0048		
	(0.0010)			(0.0099)		
$Tariff_{t-1}$				0.0278^{**}		
				(0.0102)		
$Tariff_{t-2}$				0.0168		
				(0.0099)		
$Trade_t$		0.0023^{**}			-0.0074	
		(0.0009)			(0.0297)	
$China_t$			0.0018^{**}			0.0173
			(0.0006)			(0.0319)
$China_{t-1}$						0.1515^{***}
						(0.0324)
$China_{t-2}$						0.1309^{***}
						(0.0319)
AIC	-28604.6559	-28610.2269	-28613.1693	633.4836	638.2093	610.2140
AICc	-28604.6303	-28610.2013	-28613.1437	633.7676	638.3737	610.4980
BIC	-28542.9191	-28548.4901	-28551.4325	667.4525	663.7092	644.1829
Log Likelihood	14311.3279	14314.1134	14315.5847	-308.7418	-313.1047	-297.1070
Num. obs.	7042	7042	7042	516	518	516

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 5: Summary Statistics of SPY Volatility

Model	SPY.Volatility.Mean
Full Time Mean	0.022621
First Term Mean	0.017486
Second Term Mean	0.144248