```
*Example code for SPER workshop
*The following code can be used to conduct an ARIMA anlaysis in stata
*We are going to be using the autoarima function that is now available in
*Version Stata 17

*We want to identify autocorelatioin in the residuals of the pre-interruption
*series and ensure that they are white noise in our final models.
*Can use acf and pacf to check these residuals


*Read in data
clear
set more off


*Data are extracted from CDC wonder: currently extracts raw birth counts by
month
*for the State of California for the years 2014-2021

*Change path name
import delimited "/Users/alison/Dropbox/Teaching/SPER
Workshop/California_births.csv"



*yrmonth 51 is March 2020
*Create year-month indicators in post period
gen post=.
forvalues t=1/22 {
        local lb=`t'+50
        local ub=`t'+51
        gen post_`t' =(yrmonth>=`lb' & yrmonth<`ub')
        replace post=`t' if yrmonth==`lb'
}


lab define lab_post 1 "Mar '20" 2 "Apr '20" 3 "May '20" 4 "Jun '20" 5 "Jul
'20" ///
6 "Aug '20" 7 "Sep '20" 8 "Oct '20" 9 "Nov '20" 10 "Dec '20" 11 "Jan '21" ///
12 "Feb '21" 13 "Mar '21" 14 "Apr '21" 15 "May '21" 16 "Jun '21" 17 "Jul '21"
///
18 "Aug '21" 19 "Sep '21" 20 "Oct '21" 21 "Nov '21" 22 "Dec '21"  , modify
lab values post lab_post



* Calculate residual births using mean from pre-covid period
*Month 51 is March 2020
gen birth_re = birth if yrmonth<=50
egen mean_birth = mean(birth_re)
gen birth_dem = birth/mean_birth

*Used for plotting
gen coef=.
gen se=.
```

```
* Calculate demeaned Non-ART births
gen birth_re2 = nonart_birth if yrmonth<=50
egen mean_birth2 = mean(birth_re2)
gen birth_dem2 = nonart_birth/mean_birth2

gen coef2=.
gen se2=.

*tsset our data, indicating that we are using monthly data
tsset yrmonth, monthly



*Check residuals
ac birth_dem
pac birth_dem, yw

*Use ARIMA decision tools to identify pdq parameters



*Ask stata to perform the algorithm to identify a best-fitting arima model
*on the pre-interruption time series
*This will take some time because the computer is testing multiple iterations
of
*arima parameters

*First create a training dataset
gen birth_pre=birth_dem if yrmonth<=50

*Then run algorithm. Note: this will take a while to run
*arimaauto birth_pre


*After identifying ARIMA parameters of your time series data, integrate into
an
*intterupted time series model with monthly post indicators

*For this example, our arima parameters will be arima (1,0,1) sarima
(1,1,1,12)

*ARIMA transfer function
arima birth_dem post_*, vce(oim) arima(1,0,1) sarima(1,1,1,12) technique(nr)

 forvalues t=1/22 {
       replace coef = _b[S12.post_`t'] if post_`t'==1
       replace se = _se[S12.post_`t'] if post_`t'==1
}



gen upper = coef + 1.96*se
gen lower = coef - 1.96*se


predict res, r
```

```
*Check residuals
ac res
pac res, yw



* Arima model for demeaned Non-ART births

arima birth_dem2 post_*, vce(oim) arima(1,0,1) sarima(1,1,1,12) technique(nr)
cap log close
 forvalues t=1/22 {
        replace coef2 = _b[S12.post_`t'] if post_`t'==1
        replace se2 = _se[S12.post_`t'] if post_`t'==1
}

gen upper2 = coef2 + 1.96*se2
gen lower2 = coef2 - 1.96*se2


predict res2, r


*Check residuals
ac res2
pac res2, yw



* twoway graph
replace post = post -0.2
gen post2 = post +0.4

twoway (scatter coef post, msymbol(O) msize(medium) mlcolor(navy)
mfcolor(navy)) (rspike upper lower post, lcolor(navy) lwidth(vvvthin))
(scatter coef2 post2 , msymbol(Oh) msize(medium) mlcolor(maroon) ) (rspike
upper2 lower2 post2, lcolor(maroon) lwidth(vvvthin)), yline(0) xtitle("Time
of birth", height(14)) ytitle("Effect estimate")  graphregion(color(white))
xlabel(1(1)22, valuelabel labsize(small) angle(45)) xsize(14) ysize(6)
legend(order(1 "All births" 3 "Non-ART") size(vsmall))
```