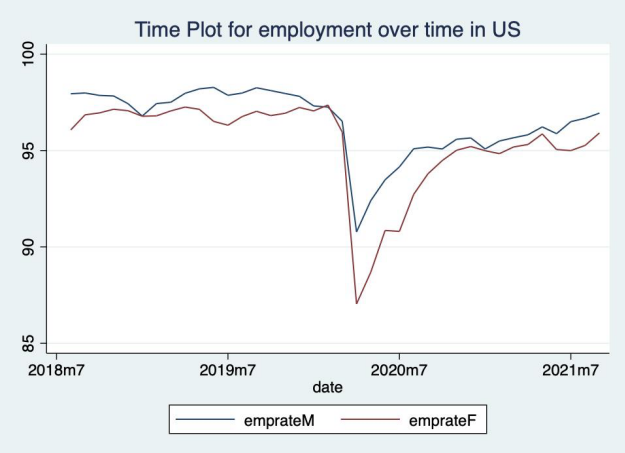
```
cd "/Users/apple/Desktop"
use "/Users/apple/Desktop/honor_CA_05.dta", clear

*************************************************************************
*Data Cleaning
*************************************************************************
*Set closure variable
gen sch_clo=0
replace sch_clo=1 if year==2021
replace sch_clo=1 if year==2020 & month==3
replace sch_clo=1 if year==2020 & month==4
replace sch_clo=1 if year==2020 & month==5
replace sch_clo=1 if year==2020 & month==6
replace sch_clo=1 if year==2020 & month==7
replace sch_clo=1 if year==2020 & month==8
replace sch_clo=1 if year==2020 & month==9
replace sch_clo=1 if year==2020 & month==10
replace sch_clo=1 if year==2020 & month==11
replace sch_clo=1 if year==2020 & month==12
*************************************************************************
*Set reopening variable; 0 for closed, 1 for partical closed, 2 for fully reopened
*Add State level static school policy
gen sch_reop=0
*Iowa
replace sch_reop=1 if state==19 & year== 2020 & month==8
replace sch_clo=0 if state==19 & year== 2020 & month==8
*************************************************************************
*Add county level static school policy
*select only CA data (I generate separate datasets for state-level and county-level)
keep if statefip==6
*select those with county level data
egen county_code = anymatch(county), values(6001 6007 6013 6017 6019 6023 6025 6029
6031 6037 6039 6041 6047 6053 6055 6059 6061 6065 6067 6071 6073 6075 6077 6079 6081
6083 6087 6089 6095 6097 6099 6107 6111 6113)
keep if county_code
*add school close data
gen sch_clo=0
replace sch_clo=1 if year==2020 & month>=3
replace sch_clo=1 if year==2021
gen sch_reopen=0
*add school reopen data
*Alameda
replace sch_reopen=1 if county==6001 & year==2021 &month >= 2 &month<=8
replace sch_reopen=2 if county==6001 & year==2021 &month ==9
```

```
replace sch_clo=0 if county==6001 & year==2021 &month >= 2 &month<=9
*****************************************************************************
*Add county dynamic school policy
*Alameda:reopen 2021 Feb
replace post_reopen_1=1 if county==6001 & year==2021 & month==3

forvalues i=1/12 {
	gen pre_reopen_`i'=0
	replace pre_reopen_`i'=1 if sch_reopen==0 & sch_reopen[_n+`i']>0
}

forvalues i=1/12 {
	gen post_reopen_`i'=0
	replace post_reopen_`i'=1 if sch_reopen>0 & sch_reopen[_n+1-`i']==0
}

*****************************************************************************
*add COVID related data
*combine vacc rate
merge m:1 year month county using "/Users/apple/Desktop/CA_vacc.dta"
drop if _merge==2
drop _merge

*combine cases
merge m:1 year month county using "/Users/apple/Desktop/CA_new_cases_pos_rate.> dta"
drop if _merge==2
drop _merge

*clean the dots in dataset
replace confirmed_cases=0 if confirmed_cases==.
replace new_cases=0 if new_cases==.
replace positive_rate=0 if positive_rate==.
replace at_least_one_dose_rate=0 if at_least_one_dose_rate==.
replace fully_vacc_rate=0 if fully_vacc_rate==.

*****************************************************************************
*Data Restrictions and Captions
*make sure look at the effect to those who are in labor force
drop if labforce==0
drop if labforce==1
drop if empstat==0
gen employed=0
replace employed=1 if empstat==1
replace employed=1 if empstat==10
```

```
replace employed=1 if empstat==12

*recode sex to be 0~male;1~female
gen female=0
replace female=1 if sex==2

*make sure look at the effect to those who have children in school
gen ch = inrange(yngch, 5, 18) | inrange(eldch, 5, 18)
drop if ch==0
*set combine year month into date variable
gen date = ym(year,month)
format date %tm

*simplify industry
gen industry=1 if ind>=170 &ind<=290
replace industry=2 if ind>=370 &ind<=490
replace industry=3 if ind==770
replace industry=4 if ind>=1070 &ind<=3990
replace industry=5 if ind>=4070 &ind<=5790
replace industry=6 if ind>=6070 &ind<=6390
replace industry=6 if ind>=570 &ind<=690
replace industry=7 if ind>=6470 &ind<=6780
replace industry=8 if ind>=6870 &ind<=7190
replace industry=9 if ind>=7270 &ind<=7790
replace industry=10 if ind>=7860 &ind<=8470
replace industry=11 if ind>=8560 &ind<=8690
replace industry=12 if ind>=8770 &ind<=9290
replace industry=13 if ind>=9370 &ind<=9590
replace industry=13 if ind==9890
drop if industry==.

*add occupation
*Management, Business, Science, and Arts Occupations
*Management, Business, and Financial Occupations
gen occupation=1 if occ>=10 & occ<=960
*Computer, Engineering, and Science Occupations
replace occupation=2 if occ>=1005 & occ<=1980
*Education, Legal, Community Service, Arts, and Media Occupations
replace occupation=3 if occ>=2001 & occ<=2970
*Healthcare Practitioners and Technical Occupations
replace occupation=4 if occ>=3000 & occ<=3550
*Service Occupations
replace occupation=5 if occ>=3601 & occ<=4665
*Sales and Office Occupations
```

```
*Sales and Related Occupations
replace occupation=6 if occ>=4700 & occ<=4965
*Office and Administrative Support Occupations
replace occupation=7 if occ>=5000 & occ<=5940
*Natural Resources, Construction, and Maintenance Occupations
*Farming, Fishing, and Forestry Occupations
replace occupation=8 if occ>=6005 & occ<=6130
*Construction and Extraction Occupations
replace occupation=9 if occ>=6200 & occ<=6950
*Installation, Maintenance, and Repair Occupations
replace occupation=10 if occ>=7000 & occ<=7640
*Production, Transportation, and Material Moving Occupations
*Production Occupations
replace occupation=11 if occ>=7700 & occ<=8990
*Transportation and Material Moving Occupations:
replace occupation=12 if occ>=9005 & occ<=9760
*Military Specific Occupations
replace occupation=13 if occ>=9800 & occ<=9830
drop if occupation==.

*add education dummy
gen education=0 if educ<=79
replace education=1 if educ>=80
drop if education==.

*add difficulty dummy
gen difficulty=0 if diffany==1
replace difficulty=1 if diffany==2
drop if difficulty==.

*add race dummy (1 as white)
gen white=race==100

keep if year==2018 & month>=8 | year>=2019

*add income use interval mean as number for family income
gen income=0 if faminc==100
replace income=5000 if faminc==210
replace income=7500 if faminc==300
replace income=10000 if faminc==430
replace income=12500 if faminc==470
replace income=15000 if faminc==500
replace income=20000 if faminc==600
replace income=25000 if faminc==710
```

```
replace income=30000 if faminc==720
replace income=35000 if faminc==730
replace income=40000 if faminc==740
replace income=50000 if faminc==820
replace income=60000 if faminc==830
replace income=75000 if faminc==841
replace income=100000 if faminc==842
replace income=150000 if faminc==843

*******************************************************************************
*Data Descriptions
*Creating time plot for employment
sort year serial
bysort year  female: gen tot=_n
bysort year  female: gen TOT=_N
bysort year  female : egen emp=sum(empstat)
gen emprate=100*emp/TOT
gen emprateM=100*emp/TOT if female==0
gen emprateF=100*emp/TOT if female==1
twoway(line emprateM year if female==0)(line emprateF year if female==1)
```




```
*******************************************************************************
*Summary table
*summarize employed
summarize employed if female==0 & year<=2019|female==0 & year==2020 & month<=2
summarize employed if female==1 & year<=2019|female==0 & year==2020 & month<=2
summarize employed if female==0 & year==2020 & month>=3 &month<=8
summarize employed if female==1 & year==2020 & month>=3 &month<=8
summarize employed if female==0 & year==2020 & month>=9|female==0 & year==2021
summarize employed if female==1 & year==2020 & month>=9|female==1 & year==2021
*summarize other variables
summarize female white education income if year<=2019|year==2020 & month<=2
summarize female white education income if year==2020 & month>=3 &month<=8
```

```
summarize female white education income if year==2020 & month>=9|year==2021
*T-test between pre-closure & closure period and closure & reopen period
ttest employed if female==0 & period!=2, by(period)
ttest employed if female==0 & period!=0, by(period)
ttest employed if female==1 & period!=2, by(period)
ttest employed if female==1 & period!=0, by(period)
**********************************************************************************
*Regression: Static Model:Logit
**State-Level
logit employed female sch_clo female_clo age new_cases fully_vacc_rate income difficulty
education white  i.statefip i.industry i.occ_time  if year<=2019 |year==2020 & month<=7
outreg2 using sta_state_hybrid.doc, replace ctitle(closure only) keep(female sch_clo female_clo
age new_cases fully_vacc_rate income difficulty education white)

logit employed female sch_reopen female_reopen age new_cases fully_vacc_rate income
difficulty education white  i.statefip i.industry i.occ_time if year==2020 & month>=4 |
year==2021
outreg2 using sta_state_hybrid.doc, append ctitle(reopen only) keep(female sch_reopen
female_reopen age new_cases fully_vacc_rate income difficulty education white)

logit employed female sch_clo female_clo sch_reopen female_reopen age new_cases
fully_vacc_rate income difficulty education white  i.statefip i.industry i.occ_time
outreg2 using sta_state_hybrid.doc, append ctitle(closure+reopen) keep(female sch_clo
female_clo sch_reopen female_reopen age new_cases fully_vacc_rate income difficulty
education white)
**********************************************************************************
*Dynamic Model:Event Study-CA County Level-linear model
reg employed pre_reopen_12 pre_reopen_11 pre_reopen_10 pre_reopen_9 pre_reopen_8
pre_reopen_7 pre_reopen_6 pre_reopen_5 pre_reopen_4 pre_reopen_3 pre_reopen_2
pre_reopen_1 post_reopen_1 post_reopen_2 post_reopen_3 post_reopen_4 post_reopen_5
post_reopen_6 post_reopen_7 post_reopen_8 post_reopen_9 post_reopen_10
post_reopen_11 post_reopen_12 age new_cases fully_vacc_rate i.industry i.date i.county if
female==0
estimates store male

reg employed pre_reopen_12 pre_reopen_11 pre_reopen_10 pre_reopen_9 pre_reopen_8
pre_reopen_7 pre_reopen_6 pre_reopen_5 pre_reopen_4 pre_reopen_3 pre_reopen_2
pre_reopen_1 post_reopen_1 post_reopen_2 post_reopen_3 post_reopen_4 post_reopen_5
post_reopen_6 post_reopen_7 post_reopen_8 post_reopen_9 post_reopen_10
post_reopen_11 post_reopen_12 age new_cases fully_vacc_rate i.industry i.date i.county if
female==1
estimates store female

*Creating event study graph
```

```
forvalues i=1/12 {
        label var pre_reopen_`i' "p`i'"
        label var post_reopen_`i' "a`i'"
}

coefplot (male,label (male)) (female,label (female)), vertical drop(_cons age new_cases
fully_vacc_rate  *.industry *.county *.date pre_reopen_12 pre_reopen_11 pre_reopen_10
pre_reopen_9 pre_reopen_8  post_reopen_8 post_reopen_9 post_reopen_10 post_reopen_11
post_reopen_12 ) recast(scatter) ciopt(color(%20)) nooffsets
```