Measuring California Unemployment 2018-2019

Econ 110: Roberto Bautista

What Are We Finding?

- **Problem:** Modern social issues such as gender inequality, racial inequity, and income disparity currently impacts our society today

- **What we are finding:** Whether California unemployment from 2018-2019 changes depending on factors such as year, gender, race, age, income, veteran status, and education

Regression Model:

Unemployed = B1+B2y2019+B3female+B4race+B5education+B6income+B7veteran

((Race = White, Black, American (Latinx), Chinese, Japanese, Other))

Veteran = Unemployed veteran or Employed Veteran

Income = Income Level

Education = GED, Undergraduate studies, Postgraduate

Data

- Used IPUMS as the source of our data
- Focused on unemployment in 2018 - 2019 in California, measured by county

Steps we took:

- Looked for data from 2018-2019 for unemployment on IPUMS
- Finding the variables and sample sizes
- Running the data from IPUMS on STATA
- Cleaning the data using Do File
- Doing 3 different regressions

STATA Do File

```
keep if year == 2018| year == 2019
tab empstat
gen unemp = 1 if empstat == 2
replace unemp = 0 if unemp ==.
keep if statefip == 6
tab sex
gen female = 1 if sex == 2
replace female = 0 if female ==.
reg unemp female
tab race
gen white = 1 if race == 1
replace white = 0 if white ==.
gen black = 1 if race == 2
replace black = 0 if black ==.
gen american = 1 if race == 3
replace american = 0 if american ==.
gen chinese = 1 if race == 4
replace chinese = 0 if chinese ==.
gen japanese = 1 if race == 5
replace japanese = 0 if japanese ==.
gen other = 1 if race >= 6
replace other = 0 if other ==.
tab vetstat
gen veteran = 1 if vetstat == 2
replace veteran = 0 if veteran ==.
gen y2019 =1 if year ==2019
replace y2019 =0 if y2019==.
reg unemp female white black american chinese japanese other  veteran age educ inctot y2019
reg unemp female white black american japanese other age educ inctot y2019
reg unemp female white black american other age educ inctot y2019
```

1st Regression

```
. reg unemp female white black american chinese japanese other  veteran age educ inctot y2019
note: japanese omitted because of collinearity
```

Source	SS	df	MS	Number of obs	=	758,908
				F(11, 758896)	=	948.92
Model	254.61903	11	23.1471846	Prob > F	=	0.0000
Residual	18511.8435	758,896	.024393123	R-squared	=	0.0136
				Adj R-squared	=	0.0136
Total	18766.4625	758,907	.024728277	Root MSE	=	.15618

unemp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
female	-.0025081	.0003666	-6.84	0.000	-.0032266	-.0017896
white	.0052075	.0020212	2.58	0.010	.0012461	.0091689
black	.0178978	.0021697	8.25	0.000	.0136453	.0221503
american	.0115985	.0027251	4.26	0.000	.0062575	.0169395
chinese	-.0000797	.0021731	-0.04	0.971	-.0043389	.0041795
japanese	0	(omitted)				
other	.0048092	.0020419	2.36	0.019	.000807	.0088113
veteran	.0009678	.000908	1.07	0.287	-.0008119	.0027475
age	-.0007521	.0000103	-72.82	0.000	-.0007723	-.0007318
educ	-.000668	.0000718	-9.30	0.000	-.0008087	-.0005272
inctot	-6.52e-09	7.56e-11	-86.32	0.000	-6.67e-09	-6.37e-09
y2019	-.002025	.0003586	-5.65	0.000	-.0027279	-.0013221
_cons	.0678418	.0021734	31.21	0.000	.0635819	.0721017

1st Regression

- In our first regression we used all of our variables from our model; female, race, veteran, age, education, inctot, and year 2019
- The Adjusted R-square was 0.0136 which is 1.36%, it shows how much variance is explained by our model
- Our F test is 948.92, meaning our variables are jointly significant
- After reviewing the regression we did notice that there were two insignificant variables; Chinese and Veteran
- We knew these two variables were insignificant because the p values were greater than 0.05, so we got rid of them and ran another regression

2nd Regression

```
. reg unemp female white black american japanese other age educ inctot y2019
```

Source	SS	df	MS		
Model	254.591321	10	25.4591321	Number of obs	= 758,908
Residual	18511.8712	758,897	.024393127	F(10, 758897)	= 1043.70
				Prob > F	= 0.0000
				R-squared	= 0.0136
				Adj R-squared	= 0.0136
Total	18766.4625	758,907	.024728277	Root MSE	= .15618

unemp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
female	-.0025869	.0003591	-7.20	0.000	-.0032907	-.0018832
white	.0053301	.0008647	6.16	0.000	.0036354	.0070248
black	.0180419	.0011681	15.44	0.000	.0157524	.0203314
american	.0117284	.0020197	5.81	0.000	.00777	.0156869
japanese	.0000994	.002173	0.05	0.964	-.0041597	.0043585
other	.0049141	.0009055	5.43	0.000	.0031393	.0066888
age	-.0007497	.0000101	-74.33	0.000	-.0007695	-.00073
educ	-.0006649	.0000718	-9.27	0.000	-.0008055	-.0005242
inctot	-6.52e-09	7.54e-11	-86.48	0.000	-6.66e-09	-6.37e-09
y2019	-.0020268	.0003586	-5.65	0.000	-.0027297	-.0013239
_cons	.0676842	.0011316	59.81	0.000	.0654663	.0699021

2nd Regression

- In our second regression we used majority of our variables as stated before; female, age, education, inctot, and year 2019
- In this regression we ran it without Chinese and Veteran
- The Adjusted R-square stayed the same at 1.36%, showing the variance in our model
- Our F test is 1,043.70, meaning our variables are jointly significant
- After getting rid of these two insignificant variables we got another insignificant variable which was Japanese

3rd Regression

```
. reg unemp female white black american other age educ inctot y2019
```

Source	SS	df	MS		
Model	254.59127	9	28.2879189	Number of obs	= 758,908
Residual	18511.8712	758,898	.024393095	F(9, 758898)	= 1159.67
				Prob > F	= 0.0000
				R-squared	= 0.0136
				Adj R-squared	= 0.0136
Total	18766.4625	758,907	.024728277	Root MSE	= .15618

unemp	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
female	-.0025868	.0003591	-7.20	0.000	-.0032906	-.0018831
white	.0053155	.0008038	6.61	0.000	.0037402	.0068909
black	.0180274	.0011241	16.04	0.000	.0158242	.0202306
american	.0117139	.0019947	5.87	0.000	.0078044	.0156234
other	.0048996	.0008486	5.77	0.000	.0032363	.0065629
age	-.0007497	.0000101	-74.39	0.000	-.0007695	-.00073
educ	-.0006648	.0000718	-9.27	0.000	-.0008055	-.0005242
inctot	-6.52e-09	7.53e-11	-86.50	0.000	-6.66e-09	-6.37e-09
y2019	-.0020268	.0003586	-5.65	0.000	-.0027297	-.001324
_cons	.0676977	.0010925	61.97	0.000	.0655565	.069839

3rd Regression

- In our third regression we used some of our variables like female, age, education, inctot, and year 2019
- In this regression we ran it without Chinese, Veteran, and Japanese
- The Adjusted R-square continued to be at 1.36%, illustrating the variance in our model
- Our F test is 1,159.67, meaning our variables are jointly significant
- After getting rid of these three insignificant variables we got all of our remaining variables to be significant, meaning the p values were all at 0.00

Coefficient

- In the 1st regression our y2019 coefficient was at -0.002025
- In the 2nd regression our y2019 coefficient was at -0.0020268
- In the 3rd regression our y2019 coefficient was at -0.0020268
- Based off our coefficients we can’t get a good measurement because it is not in percentage form, but we can see that the unemployment rate was lower in 2019 than 2018

Conclusion

- Based of our regressions we see that the variables that do not contribute towards changes in Unemployment are Chinese, Veteran, and Japanese
- Our variables such as Female, education, race and income are significant to unemployment rate and as they shift so does unemployment as seen by the decrease from 2018-2019
- Due to this, we can conclude that a person's education level, race, income level and gender have a statistically significant impact on employment

Problems Encountered

- Wanted to looked at effect of Covid19 on unemployment
- Lack of 2020 unemployment data by states and country
- Finding the right sample size/variables for Ipums?
- Getting our numbers into percentage form through STATA
- Issues/codings with STATA
    - Cleaning the data
    - Running regressions
    - Generating a dummy variable
    - Can not find the changes in unemployment rate by percentage because unemployment is a dummy variable