Problem Set 2
Due Class 5, February 12

PPPA 8022
Spring 2025

Some overall instructions

- Use a do-file (or its SAS or SPSS or R equivalent) for this work. Do not program interactively. While interactive programming may seem faster at first, inevitably you find mistakes and lose track of edits to the data – and it is slower.
- Turn in a typed up set of answers that answers the questions below. Also turn in a Stata .do file and its associated .log file or the equivalent in whichever software you use.
- Make formal tables to present your results. Do not present statistical software output.
- I have provided Stata datasets, but you should feel free to do the analysis in whatever software you prefer. For loading these data in R, I recommend the `haven` package, with `read_dta()`. If you need another format, please let me know.
- This problem set uses some large data. For the Census data, I have posted full dataset as well a smaller version; use whichever you prefer. For the CPS, we are using a random sample.
- If the question is insufficiently clear, explain the assumptions you made to reach your final estimates.
- Data are
  - Decennial Census data
    - Large: 1950 and 2010
    - Small: 1950 and 2010
    - Be in touch if you would like csv versions of these files
  - Current Population Survey (CPS)
    - Stata format
    - CSV

1. Interpreting Indicator Variables

For this problem, we'll use Decennial Census/American Community Survey data from IPUMS-USA for 1950 and 2010 (for 2010, the 1-year American Community Survey), linked above.

For purposes of **this problem set only** we will not use any survey-defined weights. This is totally wrong and you should never do it when you really analyze a dataset. We are doing it here so that 1(b) does not become extremely difficult.

The IPUMS website is https://usa.ipums.org/usa/, and it provides detailed information on the datasets and variables.

Let's examine the effect of education on wages.

(a) In a well-labeled table, report the mean, standard deviation and number of observations for income from wages (`incwage`; I have re-coded top coded values (99999) to missing) for prime age men (`sex == 1`, ages 25 to 64) in 1950 and 2010.

Answer: See Table 1 below.

(b) Use the values in the table from part (a) to calculate a t-test for whether 1950 average income from wages for prime-age men from (a) differs from the 2010 figure. Update your previous table to include the t value. Beware of missing values. Write a sentence or two to interpret your table.

Answer: See t-test in Table 1. I can reject that average wages for prime age men in 2010 are significantly different than average wages for the same group in 1950.

(c) Use a regression to do the same test as in (b). Write the regression equation you're estimating. Estimate the equation using software. Report the results in a well-labeled table.

Answer: Here is the regression equation:

$$\text{wage}_{i,t} = \alpha_0 + \alpha_1 1\{\text{year} = 2010\}_{i,t} + \epsilon_{i,t}$$

Here i denotes prime age men, and t denotes year 1950 or 2010. The expression $1\{\text{year} = 2010\}_{i,t}$ is an indicator equal to 1 if the year is 2010 and is zero is the year is 1950 (there are only two years in this dataset). If we can reject $\hat{\alpha} = 0$, then wages in 2010 are statistically different than wages in 1950.

See results in Table 2.

(d) Show how you can combine your estimated regression coefficients from (c) to yield either the 1950 or 2010 mean in the table from (a).

Answer: Mean of 1950 wages is $\hat{\alpha_0}$. Mean of 2010 wages is $\hat{\alpha_0} + \hat{\alpha_1}$.

From the table, we know that 1950 mean income is \$2,182, and 2010 mean income is \$43,244.

(All the "big" sample numbers.)

From the regression results, we know that $\hat{\alpha_0} = 2,182$, and $\hat{\alpha_1} = 41,062$ (again, the "big" sample numbers). The coefficient $\hat{\alpha_0}$ is the same as the table's mean 1950 income. Adding the two coefficients gives us 2010 mean income of $2,182 + 41,062 = 43,244$.

(e) Re-do the means in (a) using the constant 2022 dollar income (`rl_wage`). Report this income in a well-labled table.[1]

Answer: See Table 1.

(f) Suppose we would like to know whether the average husband earns higher real income from wages than the average wife.

1. Write a regression equation that we could use to test this hypothesis. Let $X_{i,t}$ represent covariates, $H_{i,t}$ to indicate 1 if the person is a husband, and $Y_{i,t}$ represent the outcome.

   Answer: The regression equation that tests whether the average husband earns higher real wages is

   $$Y_{i,t} = \beta_0 + \beta_1 H_{i,t} + \beta_2 X_{i,t} + \epsilon_{i,t}$$

2. Use Stata to estimate the regression from (f.1) using covariates age and year. Think about what sample you should use to do this, and explain what sample you choose. Report results in a well-labeled table. Make sure you only keep working age people. Here and elsewhere, your tables need only include the relevant coefficients; do not report information on all coefficients.

[1] For your information, here is how I adjust for inflation:

- Go to the Bureau of Labor Statistics (http://www.bls.gov/cpi/data.htm, and choose "all urban consumers" row and the "top picks" column.
- From the following window, choose the "US city average, All items" and choose "retrieve data," at the bottom.
- Download the data using the xls icon, making sure you're grabbing the relevant years; see the selection at top.
- Use the December inflation number for each year (this is not exactly correct, but it is sufficient for this example).
- To inflation adjust
  - re-scale the inflation adjustment so that it is 1 in 2022
  - to do this, divide the 2022 value by each year's value
  - this gives 1 in 2022, numbers > 1 in years before 2022 and numbers < 1 in years after 2022
  - this new ratio is the adjustment factor
  - multiply the adjustment factor by the values (e.g., wage) I make into constant dollars

Answer: See program for details on the sample. In brief, I kept only those who are married, and those age 25 to 65. I keep only married people so I am comparing husbands in 1950 to husbands in 2010.

See Table 3 at end, columns 1 and 5.

3. Modify the regression equation from (f.1.) to allow the relationship between being a husband and earnings to vary between 1950 and 2010.

   Answer: The regression equation that tests whether husbands' average real income from wages is also higher in 2010 is

   $$Y_{i,t} = \beta_0 + \beta_1 H_{i,t} + \beta_2 H_{i,t} * 1\{\text{year} = 2010\} + \beta_3 1\{\text{year} = 2010\} + \beta_4 X_{i,t} + \epsilon_{i,t}$$

   The key coefficients in this estimation are

   - β_1: measures the average additional wage for husbands
   - β_2: measures average additional wages for husbands in 2010
   - β_3: measure average wages in 2010

   The coefficient β_2 tests whether being a husband is associated with higher wages in 2010, relative to 1950. You need all three terms to separate out the impact over being a husband overall (β_1), and being in 2010 overall (β_3).

4. Estimate the regression from (f.3), and report the results in a well-labled table. Write a sentence to interpret your test of whether the relationship between being a husband and earnings is different in 2010 than in 1950.

   Answer: See Table 3 at end, columns 2 and 6. On average, in 1950, husbands earn roughly \$25,000 dollars more than wives. This gap increases by about \$12,000 in 2010, so that the total 2010 gap between husband and wife real income from wages is roughly \$37,000.

(g) The previous estimation included age linearly. Use the estimation for (f.1) and use a method that relaxes the linear assumption on age. Any method is fine. Report the results in a table. Write a few sentences that interpret the results, comparing with part (f).

Answer: See Table 3 at end. I relax the linearity assumption in two ways. In columns 3 and 7, I include terms for age^2, age^3 and age^4. In columns 4 and 8, I omit the linear age term and include indicator variables for each age. Coefficients are all expressed relative to average real income from wages at age 25. I don't include all these coefficients for space reasons, however. You can see from the two coefficients I do report that income from wages increases in age.

Regardless of how I specify the estimation, the main results on the relationship between being a husband, and being a husband in 2010 with real income from wages is unchanged.

2. Difference-in-difference

Now let's use the IPUMS–CPS; data are linked above. Documentation for this dataset is available at https://cps.ipums.org/cps/. For the purposes of this problem set, treat each observation with equal weight. This is entirely wrong, and you should absolutely never do such a thing if you are doing a real project. Finally, beware of top-coded data!

(a) Pretend that MI, CA, AZ, NM, MN, OH, VA, KY, WV, MO, MS, GA, IA, NH, MA and ME all adopt a policy aimed at increasing wages that takes effect in 2000. For simplicity, focus only on employed people for this entire question. Use the variable `incwage` for annual wages. Create a figure that examines the parallel pre-trend assumption.

Hints on how to create this figure:

- Sketch yourself what this graph should look like
- Then ask "what summary statistics do I need to make this graph?"
- Create the summary statistics
- Plot the summary statistics

Write a few sentences that interpret the figure.

Answer: See Figure 1 at the end. I don't see any compelling difference between the two groups pre-treatment (in the pre-2000 era). I've added gray bands that show the 95% confidence intervals for the means. This is Stata command `rarea`, and it can be a very helpful way to show a lot of information.

(b) Suppose we hypothesize that treatment is random conditional on age and race.

1. Write a regression equation to test whether the treated and untreated states have similar trends before the treatment is adopted, conditional on covariates. Look at Lecture 4 for our discussion of trends.

   Answer: These regressions are relevant for only the pre-treatment data. I tested the equality of trends in two different ways:

$$\text{incwage}_{i,t} = \beta_0 + \beta_1 \text{trend}_t + \beta_2 \text{trend}_t * \text{treatment}_{i,t} + \epsilon \quad (1)$$

$$\text{incwage}_{i,t} = \beta_0 + \beta_1 \text{time}_t + \beta_2 \text{time}_t * \text{treatment}_{i,t} + \epsilon \quad (2)$$

The variable treatment is 1 if the state is ever treated. The variable trend_t is a linear trend variable, e.g., 1960=1. 1961=2, etc., though the exact number for each year is not consequential for the slope, only the intercept. The variable time_t is a full set of year dummy variables (that is, fixed effects), with one indicator variable for each year (minus one).

Test whether treated and untreated states have differing pre-treatment trends conditional on covariates by testing whether $\beta_2 = 0$. For the first equation, this is a test on a single coefficient. For the second equation, this is a test of whether a set of coefficients, one for each time period, are all equal to zero.

2. Estimate the equation from the previous step. Report the results in a table, and write a few sentences that interpret the results of your test.

   Some advice for this estimation: Limit the sample to only pre-treatment data. Create a time trend variable. Then regress the outcome of interest on this time trend variable interacted with treatment.

   Answer: We test $H_0 : \beta_2 = 0$. For Equation 2, all we need is a t-test for whether $\beta_2 = 0$. I find a t-value $= 9.9/7.4 = 1.3$, so we cannot reject $\beta_2 = 0$.

   If instead we estimated Equation 2, we want to know whether all the β_2 are jointly 0. We do the second with an F test ($H_0 : \beta_{2,1963} = \beta_{2,1964} = \ldots = \beta_{2,1999} = 0$). We cannot reject the hypothesis that the $\text{year}_t * \text{treatment}_{i,t}$ coefficients are jointly zero.

```
.  local testvals _IyeaXtre_1963;


.  forvalues y=1964/1999
> ;
2.  local testvals ‘testvals’ = _IyeaXtre_‘y’;
3.  ;


.  test ‘testvals’ = 0;


( 1) _IyeaXtre_1963 - _IyeaXtre_1964 = 0
( 2) _IyeaXtre_1963 - _IyeaXtre_1965 = 0
( 3) _IyeaXtre_1963 - _IyeaXtre_1966 = 0
( 4) _IyeaXtre_1963 - _IyeaXtre_1967 = 0
( 5) _IyeaXtre_1963 - _IyeaXtre_1968 = 0
( 6) _IyeaXtre_1963 - _IyeaXtre_1969 = 0
( 7) _IyeaXtre_1963 - _IyeaXtre_1970 = 0
( 8) _IyeaXtre_1963 - _IyeaXtre_1971 = 0
( 9) _IyeaXtre_1963 - _IyeaXtre_1972 = 0
(10) _IyeaXtre_1963 - _IyeaXtre_1973 = 0
```

```
(11) _IyeaXtre_1963 - _IyeaXtre_1974 = 0
(12) _IyeaXtre_1963 - _IyeaXtre_1975 = 0
(13) _IyeaXtre_1963 - _IyeaXtre_1976 = 0
(14) _IyeaXtre_1963 - _IyeaXtre_1977 = 0
(15) _IyeaXtre_1963 - _IyeaXtre_1978 = 0
(16) _IyeaXtre_1963 - _IyeaXtre_1979 = 0
(17) _IyeaXtre_1963 - _IyeaXtre_1980 = 0
(18) _IyeaXtre_1963 - _IyeaXtre_1981 = 0
(19) _IyeaXtre_1963 - _IyeaXtre_1982 = 0
(20) _IyeaXtre_1963 - _IyeaXtre_1983 = 0
(21) _IyeaXtre_1963 - _IyeaXtre_1984 = 0
(22) _IyeaXtre_1963 - _IyeaXtre_1985 = 0
(23) _IyeaXtre_1963 - _IyeaXtre_1986 = 0
(24) _IyeaXtre_1963 - _IyeaXtre_1987 = 0
(25) _IyeaXtre_1963 - _IyeaXtre_1988 = 0
(26) _IyeaXtre_1963 - _IyeaXtre_1989 = 0
(27) _IyeaXtre_1963 - _IyeaXtre_1990 = 0
(28) _IyeaXtre_1963 - _IyeaXtre_1991 = 0
(29) _IyeaXtre_1963 - _IyeaXtre_1992 = 0
(30) _IyeaXtre_1963 - _IyeaXtre_1993 = 0
(31) _IyeaXtre_1963 - _IyeaXtre_1994 = 0
(32) _IyeaXtre_1963 - _IyeaXtre_1995 = 0
(33) _IyeaXtre_1963 - _IyeaXtre_1996 = 0
(34) _IyeaXtre_1963 - _IyeaXtre_1997 = 0
(35) _IyeaXtre_1963 - _IyeaXtre_1998 = 0
(36) _IyeaXtre_1963 - _IyeaXtre_1999 = 0
(37) _IyeaXtre_1963 = 0

F( 37,230832) = 0.71
Prob > F = 0.9093
```

These regressions each use 231,066 observations – less than the full dataset, since they omit data before 2000.

(c) Do a summary statistics version of a difference-in-difference estimate (no covariates for this question). This means find the means for the treated and untreated groups, both before and after treatment. Create a table with these means and the relevant standard errors from which you can calculate the single and double differences (don't worry about calculating the errors for these differences).

Answer: See summary statistics in Table 4 below. Income from wages declined by about

$882 in the treated states, relative to the untreated ones, after the treatment.

(d) Do a difference-in-difference regression that parallels the summary statistics in part (c), meaning that it has no covariates.

1. Write the estimating equation you use.

   Answer: To do the regression, I estimate the following equation:

   $$\text{incwage}_{ist} = \beta_0 + \beta_1 \text{after}_t + \beta_2 \text{treated}_{is} + \beta_3 \text{treated*after}_{ist} + \epsilon_{ist}$$

2. Estimate the regression and report the results in a well-labeled table. Include the constant.

   Answer: See results in Table 5

3. Show that you can combine the regression coefficients from the table you just created to add up to one of the sample means from part (c). If you don't get the same result as in (c), you are doing something wrong. The key to getting the diff-in-diff regression and the summary stats to match is to use the **exact** same sample.

   Answer: The constant in Table 5 is the mean of the untreated group "before" in Table 4. The untreated-after mean is the constant value plus the "after" coefficient of 23,916. The treated-before group mean is the Table 5 constant plus the treatment indicator. And the treated-after mean is the constant plus the treatment indicator plus the after indicator plus the treatment-after interaction.

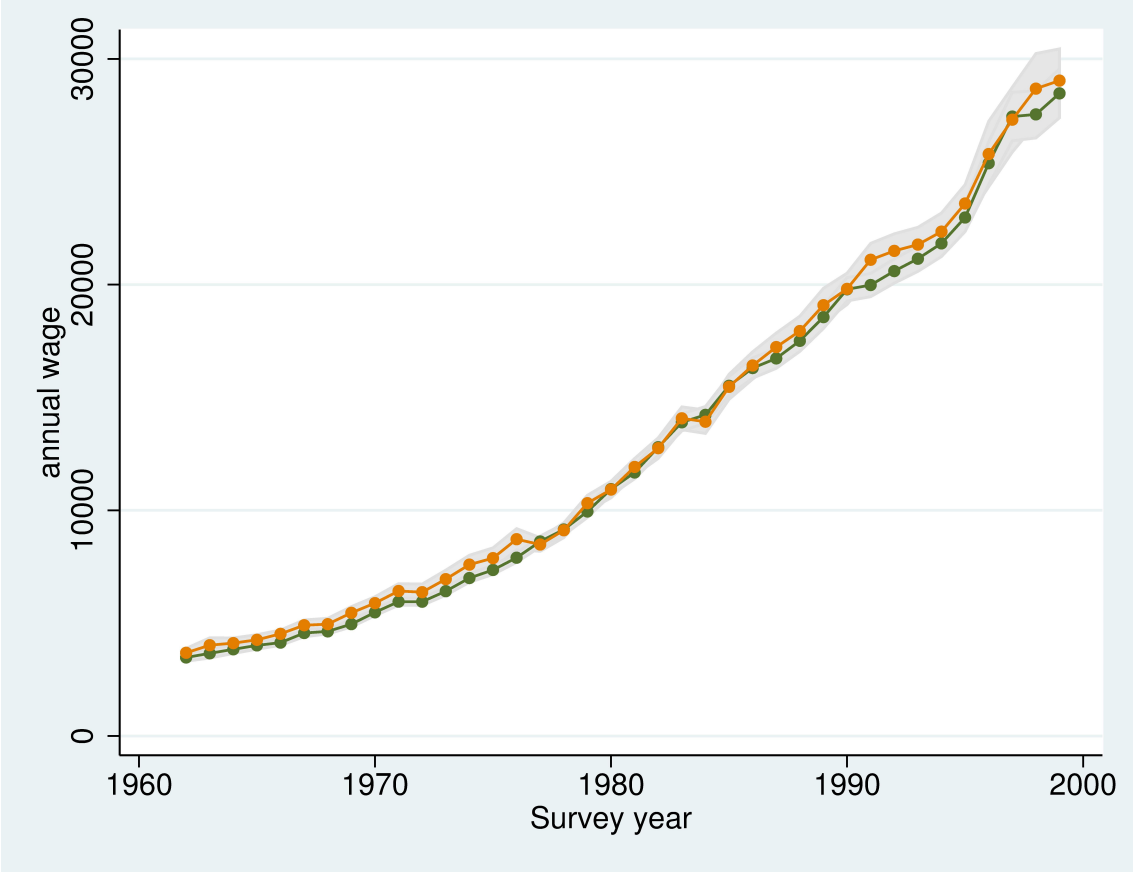

Figure 1: Parallel Pre-Trend Assumption

Table 1: Comparison of Means

	Nominal Wage Income			Real Wage Income		
	1950	2010	t test, difference	1950	2010	t test, difference
	(1)	(2)	(3)	(4)	(5)	(6)
Small Sample						
mean	2,163.3	43,405.5	201.4	25,892.8	58,857.8	93.2
std error of the mean	18.4	204.0		220.2	276.6	
observations	11,334	78,423		11,334	78,423	
Big Sample						
mean	2,182.2	43,243.8	638.4	26,118.7	58,638.6	291.7
std error of the mean	5.8	64.1		69.9	86.9	
observations	111,680	787,469		111,680	787,469	

Table 2: Comparison via Regression

	Dependent Variable is Income from Wages	
	Sample is	
	Small	Big
	(1)	(2)
ear = 2010}	41,242	41,062
	(536.61)	(170.11)
Constant	2,163	2,182
	(501.59)	(159.19)
R-squared	0.062	0.061
Observations	89,757	899,149

Table 3: Comparing Husbands' Wages

	Dependent variable is Income from Wages in 2022 Dollars							
	Sample is							
	Small				Big			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1{person is male}	35210.2***	24727.8***	24117.0***	24114.0***	34948.3***	24831.0***	24398.3***	24399.3***
	(384.3)	(991.4)	(982.2)	(982.3)	(120.9)	(313.5)	(310.8)	(310.8)
Age	-285.9***	-284.2***	15590.6**		-266.4***	-264.4***	16305.9***	
	(17.8)	(17.7)	(5315.5)		(5.6)	(5.6)	(1681.7)	
1{year = 2010}	39343.7***	33199.4***	31767.4***	31752.8***	39134.2***	33259.6***	31928.3***	31940.0***
	(545.0)	(764.0)	(757.5)	(757.8)	(172.1)	(240.5)	(238.6)	(238.7)
1{year = 2010} * 1{person is male}		12329.2***	12793.8***	12808.4***		11878.1***	12239.5***	12240.9***
		(1075.0)	(1064.9)	(1065.1)		(339.6)	(336.7)	(336.7)
Age^2			-498.5**				-522.9***	
			(184.90)				(58.40)	
Age^3			8.0**				8.3***	
			(2.80)				(0.90)	
Age^4			-0.1***				-0.1***	
			(0)				(0)	
1{age = 26}				807.7				1720.0*
				(2411.80)				(772.50)
1{age = 27}				3903.3				3305.4***
				(2338.90)				(745.50)
… and many more I omit for space, through age 65								
Constant	10499.2***	15661.5***	-192508.8***	-8780.0***	9837.4***	14758.2***	-199265.3***	-8687.7***
	(910.9)	(1015.6)	(55715.2)	(1835.8)	(286.9)	(319.4)	(17641.9)	(592.8)
R-squared	0.104	0.105	0.122	0.122	0.104	0.105	0.12	0.12
Obs	115,171	115,171	115,171	115,171	1,152,661	1,152,661	1,152,661	1,152,661

Table 4: Answer for 2(c)

			Differences	
	Before	After	Single	Double
	(1)	(2)	(3)	(4)
Treated				
mean	14910	37944	23034	
sd	68	225		
obs	73884	45534		
Untreated				
mean	13662	37578	23916	-882
sd	45	163		
obs	157170	87780		

Table 5: Regressions, Problem 2

	DD
	(1)
1{treatment}	1247.9***
	(145.3)
1{after}	23916.1***
	(137.3)
1{treatment}*1{after}	-881.1***
	(237.7)
Constant	13661.6***
	(82.2)
Observations	364368

```
# delimit;

*******************************************************************************

this is the stata file that makes the answers for problem set 2

january 28, 2014
january 30, 2014
january 31, 2014
february 6, 2014
february 7, 2014
february 10, 2014 ** fixing prime age male to be males
february 18, 2014 ** fixing inflation correction
january 31, 2018  ** adding an initial table
february 1, 2018  ** checking final stuff
january 30, 2019  ** updating for new year
february 7, 2019  ** fixing frans errors
september 19, 2023 ** building in inflation adjustment so students dont have to do it
september 20, 2023 ** still updating
october 5, 2023    ** fixing summary stats table

ps2v15.do

*******************************************************************************;

capture log close;
log using ps2v15.log, replace;

*******************************************************************************

****** 1. ipums-usa data ***************************************************

*******************************************************************************;

******* A. get ready ********************************************************;

*** A.1. start up things ***;

clear all;
pause on;
set more off;

* set todays date;
* this makes a local macro with todays date *;
adopath ++ /home/lfbrooks/home/bleah/ado;
dateo;

*** A.2. load data ***;

* switch for which sample we use *;
local sample big;
*local sample small;

* load the big sample *;
if "`sample'" == "big"
  {;

  * 1950 *;
  * program that creates these data is
    /groups/brooksgrp/pppa6022/2023/problem_sets/ps2/load1950v02.do
    old:
/home/lfbrooks/home/bleah/pppa6022/2014/stataprg/problem_set_1/load1950v01.do;
```

```
  use
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/data_for_students/20230920_ipums_195
0;

  * 2010 *;
  * program that creates these data is
    /groups/brooksgrp/pppa6022/2023/problem_sets/ps2/load2010v02.do
    old
/home/lfbrooks/home/bleah/pppa6022/2014/stataprg/problem_set_1/load2010v01.do;
  append using
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/data_for_students/20230920_ipums_201
0;

  * code making the smaller sample *;
  * take a smaller random sample *;
  *gen double rand_samp = runiform();
  *keep if rand_samp > 0.9;

  * save this for students who need a small dataset *;
  *save
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/data_for_students/${date}_small_ipum
s, replace;
  };

* load the small sample *;
if "`sample'" == "small"
  {;
  * created in step above *;
  use
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/data_for_students/${date}_small_ipum
s;
  };

***** Q 1(a) and (b) and (e): mean wages in 1950 and 2010 *********;

**** prepare data ****;

*** marker for prime-age males ***;
gen prime_age_male = 0;
replace prime_age_male = 1 if (age >= 25 & age < 65) & sex == 1;

* make a 2010 dummy *;
gen y2010 = 0;
replace y2010 = 1 if year == 2010;

**** take means ****;

* save the data for later use *;
preserve;

*** make a small table with mean and sd of mean of wage *;
sort year prime_age_male;
collapse (mean) incwage=incwage (sem) incwage_se=incwage (count) inc_obs=incwage
         (mean) rl_wage=rl_wage (sem) rl_wage_se=rl_wage, by(year prime_age_male);

* keep only obs of interest *;
keep if prime_age_male == 1;

* make the dataset wide so I can do t-tests *;
* this next line is a programming trick to get the reshape to behave *;
gen marker = 1;
reshape wide incwage incwage_se rl_wage rl_wage_se inc_obs, i(marker) j(year);
```

```
* make the t values *;
gen t_nominal = abs((incwage2010-
incwage1950)/sqrt(incwage_se2010^2+incwage_se1950^2));
gen t_real = abs((rl_wage2010-rl_wage1950)/sqrt(rl_wage_se2010^2+rl_wage_se1950^2));

* output the dataset *;
export delimited
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/${date}_q1abe_`sample'.txt,
replace;


***** Q 1(c) and (d): regression to repeat this test *******;

* get data back *;
restore;

* make a regression that looks at average wages for men *;
regress incwage y2010 if prime_age_male == 1;
estimates store r1;
estout r1 using

/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/${date}_q1cd_`sample'.txt,
           replace label varwidth(12)
       cells(b(fmt(%9.3f)) se(par fmt(%9.3f)))
       stats(r2 N,
         fmt(%9.3f %9.0g )
         labels("R-squared" "Obs"  ));
estimates clear;

****** Q 1(f) and (g): husband vs wife wages
**************************************;

* subset to married couples *;

* keep if married, spouse present *;
keep if marst == 1;
* keep if head of household or spouse *;
keep if relate == 1 | relate == 2;

* make male dummy *;
gen male = 0;
replace male = 1 if sex == 1;

* make 2010*male *;
gen y2010_male = y2010*male;

* keep only 65 or younger *;
keep if age <= 65 & age >= 25;

* make age squared, cubed, to the fourth power *;
gen age2 = age*age;
gen age3 = age2*age;
gen age4 = age3*age;

* make a hh serial number by year (thanks, Drew! for catching this) *;
sort year;
by year: summ serial;
gen serialyr = 1950*1000000+serial if year == 1950;
replace serialyr = 2010*1000000+serial if year == 2010;

* check this sample *;
summ age male incwage;
```

```
*** regressions ***;

*** 1(f) ***;

* estimate wages as a function of age, year, and being the husband *;
regress rl_wage age male y2010;
eststo c1;
* re-estimate, allowing the main effect to vary by year*;
regress rl_wage age male y2010 y2010_male;
eststo c2;

*** 1(g): different treatments of age ***;

* parametric non-linear effects of age *;
regress rl_wage age age2 age3 age4 male y2010 y2010_male;
eststo c3;
* non-parametric non-linear effects of age *;
xi: regress rl_wage i.age male y2010 y2010_male;
eststo c4;

*** output results ***;
estout *
  using
"/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/${date}_q1fg_`sample'.txt",
  replace
  varwidth(12) varlabels(_cons Constant)
  cells(b(star fmt(%12.1f)) se(par fmt(%12.1f)))
  stats(r2 N, fmt(%9.3f %9.0g %9.3f) labels("R-squared" "Obs"));


*******************************************************************************

****** 2. ipums-cps data ***************************************************

*******************************************************************************;

****** A. prep stuff *******************************************************;

clear all;

* set todays date;
adopath ++ /home/lfbrooks/home/bleah/ado;
dateo;


****** B. load data ********************************************************;

* switch for which sample we use *;
local sample big;
*local sample small;

* load the big sample *;
if "`sample'" == "big"
  {;

  * load cps data, keeping only variables that are of interest *;
  use
/groups/brooksgrp/containerization/current_population_survey/input_201401/20140107_cp
sinput;

  * take a smaller random sample *;
  gen double rand_samp = runiform();
  keep if rand_samp > 0.9;
```

```
  * save this as a junk dataset to load *;
  save /groups/brooksgrp/junk/ipumscps, replace;
  };

if "`sample'" == "small"
  {;
  * bring in the smaller data *;
  use /groups/brooksgrp/junk/ipumscps;
  };



****** C. set up data ********************************************************;

** fix topcodes **;
* not in universe *;
replace incwage = . if incwage == 9999999;
* missing *;
replace incwage = . if incwage == 9999998;
* top-coded -- starts in 2011.. i think at 1,000,000 *;
replace incwage = 1000000 if incwage == 9999997;

* marker for employed guys *;
gen emp = 0;
replace emp = 1 if empstat == 10 | empstat == 12 | empstat ==13;

** make fake treatment **;
* make a list of all treated states *;
local trts 26 6 4 35 27 39 51 21 54 29 28 13 19 33 25 23;
* make a treatment variable *;
gen treat=0;
* fix it for the affected states *;
foreach j in
  `trts'
  {;
  replace treat=1 if statefip == `j';
  label variable treat "1 if ever in a treated state";
  };

* make a linear trend *;
gen trend = year - 1979;

* make an after variable *;
gen after=0;
replace after=1 if year >= 2000;

* make a trend*treatment interaction *;
gen trend_treat = 0;
replace trend_treat = trend * treat * after;
gen trend_treat_before = 0;
replace trend_treat_before = trend if treat == 1;


* make a treated*after variable *;
gen treat_after = treat*after;
label variable treat_after "1 if after the treatment in a treated state";

* make a treated*after*male variable *;
gen male = 0;
replace male = 1 if sex==1;
label variable male "1 if male (from sex)";
gen treat_after_male = treat*after*male;
```

```
label variable treat_after_male "1 if after the treatment in a treated state and obs
is male";

* make other interactions that we need *;
gen male_after = male*after;
gen male_treat = male*treat;


****** C. no c now *****;

****** D. pictures ************************************************************;

* keep the data around for later *;
preserve;

* command for making graphs i like *;
graph set eps orientation landscape;

*** D.1. set up data for the picture ***;

* get annual mean wages for treated and untreated states, and also the std error *;
sort year treat;
collapse (mean) incwage (sem) incwage_sd = incwage if emp == 1 & year < 2000, by(year
treat);

* reshape to make it wider *;
reshape wide incwage incwage_sd, i(year) j(treat);

* calculate confidence bands *;
foreach j in
  0 1
  {;
  gen cihi`j' = incwage`j' + 1.96*incwage_sd`j';
  label variable cihi`j' "upper confidence bound for incwage when treatment = `j'";
  gen cilo`j' = incwage`j' - 1.96*incwage_sd`j';
  label variable cilo`j' "lower confidence bound for incwage when treatment = `j'";
  };

* check this *;
list;

*** D.2. make the average wage over time picture ***;

graph twoway
  (rarea cilo1 cihi1 year, bcolor(gs14))
  (rarea cilo0 cihi0 year, bcolor(gs14))
  (connected incwage0 year, msymbol(o))
  (connected incwage1 year, msymbol(o)),
  ytitle("annual wage")
  legend(off)
  xsize(11) ysize(8.5);
graph export "output/${date}_mean_wage_wandwo_treatment_`sample'.eps", replace;



****** E. regressions and summary stats
*********************************************************;

* bring back the data *;
restore;

*** E.1. test the parallel pre-trend assumption with a regression (2(a)) ***;
```

```
* could have a linear trend, and an interaction term for the treated guys *;
xi: regress incwage trend trend_treat_before i.age i.race i.statefip if emp == 1 &
year < 2000;
eststo t1;

* could test whether year effects are equal before *;
xi: regress incwage i.year*treat i.race i.age i.statefip if emp == 1 & year < 2000;
local testvals _IyeaXtre_1963;
forvalues y=1964/1999
  {;
  local testvals `testvals' = _IyeaXtre_`y';
  };
test `testvals' = 0;
eststo t2;

*** E.2. diff-in-diff regression *****;

* regression to paralell diff in diff summary stats *;
regress incwage treat_after treat after if emp == 1;
eststo r0;

* make a marker for the regression sample *;
gen r0emp = e(sample);


*** E.3. summary tables for diff-in-diff ***;

* save the data for later *;
preserve;

sort treat after;
collapse (mean) incwage (sem) incwage_sd = incwage (count) incwage_obs=incwage
  if r0emp == 1, by(treat after);

* output to excel *;
outsheet using output/${date}_prob2c_ttests_`sample'.txt, replace;


*** E.4. output regression results ***;
estout *
  using "output/${date}_prob2_regs_`sample'.txt",
  replace
  varwidth(12) varlabels(_cons Constant)
  cells(b(star fmt(%12.1f)) se(par fmt(%12.1f)))
  stats(r2 N, fmt(%9.3f %9.0g %9.3f) labels("R-squared" "Observations"));


log close;
```

```
-------------------------------------------------------------------------------
-----------------------------------------
      name:  <unnamed>
       log:  /groups/brooksgrp/pppa6022/2023_fall/problem_sets/ps2/ps2v15.log
  log type:  text
 opened on:   6 Oct 2023, 00:02:33

. ********************************************************************************
>
> ****** 1. ipums-usa data ***************************************************
>
>
********************************************************************************;
. ******* A. get ready
*************************************************************;
. *** A.1. start up things ***;
. clear all;

. pause on;

. set more off;

. * set todays date;
. * this makes a local macro with todays date *;
. adopath ++ /home/lfbrooks/home/bleah/ado;
  [1]              "/home/lfbrooks/home/bleah/ado"
  [2]  (BASE)      "/usr/local/stata16/ado/base/"
  [3]  (SITE)      "/usr/local/ado/"
  [4]              "."
  [5]  (PERSONAL)  "~/ado/personal/"
  [6]  (PLUS)      "~/ado/plus/"
  [7]  (OLDPLACE)  "~/ado/"

. dateo;

. *** A.2. load data ***;
. * switch for which sample we use *;
. local sample big;

. *local sample small;
. * load the big sample *;
. if "`sample'" == "big"
>   {;
.   * 1950 *;
.   * program that creates these data is
>     /groups/brooksgrp/pppa6022/2023/problem_sets/ps2/load1950v02.do
>     old:
/home/lfbrooks/home/bleah/pppa6022/2014/stataprg/problem_set_1/load1950v01.do;
.   use
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/data_for_students/20230920_ipums_195
0;
.   * 2010 *;
.   * program that creates these data is
>     /groups/brooksgrp/pppa6022/2023/problem_sets/ps2/load2010v02.do
>     old
/home/lfbrooks/home/bleah/pppa6022/2014/stataprg/problem_set_1/load2010v01.do;
.   append using
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/data_for_students/20230920_ipums_201
0;
(label poverty_lbl already defined)
(label incother_lbl already defined)
(label incwage_lbl already defined)
(label inctot_lbl already defined)
```

```
(label wkswork2_lbl already defined)
(label classwkrd_lbl already defined)
(label classwkr_lbl already defined)
(label ind1990_lbl already defined)
(label ind1950_lbl already defined)
(label ind_lbl already defined)
(label occ1990_lbl already defined)
(label occ1950_lbl already defined)
(label occ_lbl already defined)
(label labforce_lbl already defined)
(label empstatd_lbl already defined)
(label empstat_lbl already defined)
(label educd_lbl already defined)
(label educ_lbl already defined)
(label school_lbl already defined)
(label racesingd_lbl already defined)
(label racesing_lbl already defined)
(label hispand_lbl already defined)
(label hispan_lbl already defined)
(label citizen_lbl already defined)
(label bpld_lbl already defined)
(label bpl_lbl already defined)
(label raced_lbl already defined)
(label race_lbl already defined)
(label marrno_lbl already defined)
(label birthyr_lbl already defined)
(label birthqtr_lbl already defined)
(label marst_lbl already defined)
(label sex_lbl already defined)
(label age_lbl already defined)
(label related_lbl already defined)
(label relate_lbl already defined)
(label farm_lbl already defined)
(label gq_lbl already defined)
(label cntry_lbl already defined)
(label appald_lbl already defined)
(label appal_lbl already defined)
(label citypop_lbl already defined)
(label city_lbl already defined)
(label metaread_lbl already defined)
(label metarea_lbl already defined)
(label metro_lbl already defined)
(label county_lbl already defined)
(label statefip_lbl already defined)
(label stateicp_lbl already defined)
(label region_lbl already defined)
(label hhtype_lbl already defined)
(label subsamp_lbl already defined)
(label numprec_lbl already defined)
(label year_lbl already defined)
.       * code making the smaller sample *;
.   * take a smaller random sample *;
.   *gen double rand_samp = runiform();
.   *keep if rand_samp > 0.9;
.   * save this for students who need a small dataset *;
.   *save
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/data_for_students/${date}_small_ipum
s, replace;
.    };

. * load the small sample *;
. if "`sample'" == "small"
>   {;
```

```
.   * created in step above *;
.   use
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/data_for_students/${date}_small_ipum
s;
.   };

. ***** Q 1(a) and (b) and (e): mean wages in 1950 and 2010 *********;
. **** prepare data ****;
. *** marker for prime-age males ***;
. gen prime_age_male = 0;

. replace prime_age_male = 1 if (age >= 25 & age < 65) & sex == 1;
(1,213,439 real changes made)

. * make a 2010 dummy *;
. gen y2010 = 0;

. replace y2010 = 1 if year == 2010;
(3,061,692 real changes made)

. **** take means ****;
. * save the data for later use *;
. preserve;

. *** make a small table with mean and sd of mean of wage *;
. sort year prime_age_male;

. collapse (mean) incwage=incwage (sem) incwage_se=incwage (count) inc_obs=incwage
>          (mean) rl_wage=rl_wage (sem) rl_wage_se=rl_wage, by(year prime_age_male);

. * keep only obs of interest *;
. keep if prime_age_male == 1;
(2 observations deleted)

. * make the dataset wide so I can do t-tests *;
. * this next line is a programming trick to get the reshape to behave *;
. gen marker = 1;

. reshape wide incwage incwage_se rl_wage rl_wage_se inc_obs, i(marker) j(year);
(note: j = 1950 2010)

Data                               long   ->   wide
-----------------------------------------------------------------------------
Number of obs.                        2   ->       1
Number of variables                   8   ->      12
j variable (2 values)              year   ->   (dropped)
xij variables:
                                incwage   ->   incwage1950 incwage2010
                              incwage_se   ->   incwage_se1950 incwage_se2010
                                 rl_wage   ->   rl_wage1950 rl_wage2010
                              rl_wage_se   ->   rl_wage_se1950 rl_wage_se2010
                                 inc_obs   ->   inc_obs1950 inc_obs2010
-----------------------------------------------------------------------------

. * make the t values *;
. gen t_nominal = abs((incwage2010-
incwage1950)/sqrt(incwage_se2010^2+incwage_se1950^2));

. gen t_real = abs((rl_wage2010-
rl_wage1950)/sqrt(rl_wage_se2010^2+rl_wage_se1950^2));

. * output the dataset *;
```

```
. export delimited
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/${date}_q1abe_`sample'.txt,
replace;
(note: file
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/20231006_q1abe_big.txt not
found)
file /groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/20231006_q1abe_big.txt
saved

. ***** Q 1(c) and (d): regression to repeat this test *******;
. * get data back *;
. restore;

. * make a regression that looks at average wages for men *;
. regress incwage y2010 if prime_age_male == 1;

      Source |       SS           df       MS      Number of obs   =   899,149
-------------+----------------------------------   F(1, 899147)    =  58269.18
       Model |  1.6491e+14         1  1.6491e+14   Prob > F        =    0.0000
    Residual |  2.5447e+15   899,147  2.8302e+09   R-squared       =    0.0609
-------------+----------------------------------   Adj R-squared   =    0.0609
       Total |  2.7096e+15   899,148  3.0136e+09   Root MSE        =     53199

------------------------------------------------------------------------------
     incwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       y2010 |    41061.6   170.1047   241.39   0.000      40728.2       41395
       _cons |   2182.192   159.1906    13.71   0.000     1870.184      2494.2
------------------------------------------------------------------------------

. estimates store r1;

. estout r1 using
>
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/${date}_q1cd_`sample'.txt,
>            replace label varwidth(12)
>        cells(b(fmt(%9.3f)) se(par fmt(%9.3f)))
>        stats(r2 N,
>          fmt(%9.3f %9.0g )
>          labels("R-squared" "Obs"  ));
(note: file
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/20231006_q1cd_big.txt not
found)
(output written to
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/20231006_q1cd_big.txt)

. estimates clear;

. ****** Q 1(f) and (g): husband vs wife wages
**************************************;
. * subset to married couples *;
. * keep if married, spouse present *;
. keep if marst == 1;
(2,899,210 observations deleted)

. * keep if head of household or spouse *;
. keep if relate == 1 | relate == 2;
(93,274 observations deleted)

. * make male dummy *;
. gen male = 0;

. replace male = 1 if sex == 1;
```

```
(995,703 real changes made)

. * make 2010*male *;
. gen y2010_male = y2010*male;

. * keep only 65 or younger *;
. keep if age <= 65 & age >= 25;
(350,905 observations deleted)

. * make age squared, cubed, to the fourth power *;
. gen age2 = age*age;

. gen age3 = age2*age;

. gen age4 = age3*age;

. * make a hh serial number by year (thanks, Drew! for catching this) *;
. sort year;

. by year: summ serial;

-------------------------------------------------------------------------------
-> year = 1950

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      serial |    658,814    230386.6    133207.7          1     461130

-------------------------------------------------------------------------------
-> year = 2010

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      serial |    981,687    707365.2    405925.4          2    1397788


. gen serialyr = 1950*1000000+serial if year == 1950;
(981,687 missing values generated)

. replace serialyr = 2010*1000000+serial if year == 2010;
(981,687 real changes made)

. * check this sample *;
. summ age male incwage;

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
         age |  1,640,501    44.79893    11.02515         25         65
        male |  1,640,501    .4946099    .4999711          0          1
     incwage |  1,152,661     33762.2    50969.28          0     569000

. *** regressions ***;
. *** 1(f) ***;
. * estimate wages as a function of age, year, and being the husband *;
. regress rl_wage age male y2010;

      Source |       SS           df       MS      Number of obs   = 1,152,661
-------------+----------------------------------   F(3, 1152657)   =  44520.76
       Model |  5.6123e+14         3  1.8708e+14   Prob > F        =    0.0000
    Residual |  4.8434e+15 1,152,657  4.2020e+09   R-squared       =    0.1038
-------------+----------------------------------   Adj R-squared   =    0.1038
```

```
       Total |  5.4047e+15 1,152,660  4.6889e+09   Root MSE        =     64823

------------------------------------------------------------------------------
     rl_wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -266.4016   5.587092   -47.68   0.000    -277.3522   -255.4511
        male |   34948.33   120.9124   289.04   0.000     34711.35    35185.32
       y2010 |   39134.22   172.1462   227.33   0.000     38796.82    39471.63
       _cons |   9837.429   286.8615    34.29   0.000      9275.19    10399.67
------------------------------------------------------------------------------

. eststo c1;

. * re-estimate, allowing the main effect to vary by year*;
. regress rl_wage age male y2010 y2010_male;

      Source |       SS           df       MS      Number of obs   = 1,152,661
-------------+----------------------------------   F(4, 1152656)   =  33731.79
       Model |  5.6636e+14         4  1.4159e+14   Prob > F        =    0.0000
    Residual |  4.8383e+15 1,152,656  4.1975e+09   R-squared       =    0.1048
-------------+----------------------------------   Adj R-squared   =    0.1048
       Total |  5.4047e+15 1,152,660  4.6889e+09   Root MSE        =     64788

------------------------------------------------------------------------------
     rl_wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |  -264.3538   5.584439   -47.34   0.000    -275.2992   -253.4085
        male |   24830.99   313.5032    79.20   0.000     24216.54    25445.45
       y2010 |   33259.56   240.4502   138.32   0.000     32788.28    33730.83
  y2010_male |    11878.1   339.6184    34.97   0.000     11212.46    12543.74
       _cons |   14758.23   319.3706    46.21   0.000     14132.28    15384.19
------------------------------------------------------------------------------

. eststo c2;

. *** 1(g): different treatments of age ***;
. * parametric non-linear effects of age *;
. regress rl_wage age age2 age3 age4 male y2010 y2010_male;

      Source |       SS           df       MS      Number of obs   = 1,152,661
-------------+----------------------------------   F(7, 1152653)   =  22476.75
       Model |  6.4913e+14         7  9.2733e+13   Prob > F        =    0.0000
    Residual |  4.7555e+15 1,152,653  4.1257e+09   R-squared       =    0.1201
-------------+----------------------------------   Adj R-squared   =    0.1201
       Total |  5.4047e+15 1,152,660  4.6889e+09   Root MSE        =     64232

------------------------------------------------------------------------------
     rl_wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   16305.94   1681.676     9.70   0.000     13009.91    19601.97
        age2 |  -522.9258   58.43653    -8.95   0.000    -637.4595   -408.3922
        age3 |   8.277334   .8790426     9.42   0.000      6.55444    10.00023
        age4 |  -.0531857   .0048396   -10.99   0.000    -.0626711   -.0437003
        male |   24398.34   310.8482    78.49   0.000     23789.09     25007.6
       y2010 |   31928.34   238.6259   133.80   0.000     31460.64    32396.03
  y2010_male |    12239.5   336.7175    36.35   0.000     11579.54    12899.45
       _cons |  -199265.3   17641.92   -11.29   0.000    -233842.9   -164687.7
------------------------------------------------------------------------------

. eststo c3;

. * non-parametric non-linear effects of age *;
. xi: regress rl_wage i.age male y2010 y2010_male;
```

```
i.age             _Iage_25-65         (naturally coded; _Iage_25 omitted)

      Source |       SS           df       MS      Number of obs   = 1,152,661
-------------+----------------------------------   F(43, 1152617)  =   3662.49
       Model |  6.4969e+14         43  1.5109e+13   Prob > F        =    0.0000
    Residual |  4.7550e+15 1,152,617  4.1254e+09   R-squared       =    0.1202
-------------+----------------------------------   Adj R-squared   =    0.1202
       Total |  5.4047e+15 1,152,660  4.6889e+09   Root MSE        =     64229

------------------------------------------------------------------------------
     rl_wage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    _Iage_26 |   1720.044   772.5289     2.23   0.026      205.913    3234.174
    _Iage_27 |   3305.396   745.4848     4.43   0.000     1844.271    4766.521
    _Iage_28 |    4836.02   728.1945     6.64   0.000     3408.784    6263.257
    _Iage_29 |   6265.913   716.3788     8.75   0.000     4861.835    7669.991
    _Iage_30 |   7359.357   702.8467    10.47   0.000     5981.802    8736.913
    _Iage_31 |   8656.093   704.9488    12.28   0.000     7274.417    10037.77
    _Iage_32 |   10949.83   698.8752    15.67   0.000     9580.054     12319.6
    _Iage_33 |   12337.75    696.444    17.72   0.000     10972.74    13702.76
    _Iage_34 |   12526.32   692.7503    18.08   0.000     11168.55    13884.08
    _Iage_35 |   15145.67   688.0038    22.01   0.000     13797.21    16494.13
    _Iage_36 |   16084.15   689.1867    23.34   0.000     14733.36    17434.93
    _Iage_37 |   17241.27   685.7402    25.14   0.000     15897.24     18585.3
    _Iage_38 |   17679.81   679.5817    26.02   0.000     16347.86    19011.77
    _Iage_39 |   19199.57   673.6315    28.50   0.000     17879.28    20519.87
    _Iage_40 |    18657.8   669.9952    27.85   0.000     17344.63    19970.97
    _Iage_41 |   19765.08    677.894    29.16   0.000     18436.43    21093.73
    _Iage_42 |   19538.87   676.0785    28.90   0.000     18213.78    20863.97
    _Iage_43 |    19826.9   676.7869    29.30   0.000     18500.42    21153.38
    _Iage_44 |   20521.08   676.4977    30.33   0.000     19195.17    21846.99
    _Iage_45 |   20200.42   668.0053    30.24   0.000     18891.15    21509.69
    _Iage_46 |   20810.42   667.9476    31.16   0.000     19501.27    22119.58
    _Iage_47 |   20059.17   666.8232    30.08   0.000     18752.22    21366.12
    _Iage_48 |   19922.71   667.2526    29.86   0.000     18614.92     21230.5
    _Iage_49 |    19756.6   664.6511    29.72   0.000      18453.9    21059.29
    _Iage_50 |   17320.81   660.5317    26.22   0.000     16026.19    18615.43
    _Iage_51 |   19037.75   667.0007    28.54   0.000     17730.45    20345.04
    _Iage_52 |   18589.11    664.927    27.96   0.000     17285.88    19892.35
    _Iage_53 |    17578.1   665.5493    26.41   0.000     16273.64    18882.55
    _Iage_54 |   16555.55   667.4148    24.81   0.000     15247.44    17863.66
    _Iage_55 |   14816.46   666.1607    22.24   0.000     13510.81    16122.11
    _Iage_56 |   13185.23   670.2589    19.67   0.000     11871.54    14498.91
    _Iage_57 |   11801.56   671.2427    17.58   0.000     10485.95    13117.17
    _Iage_58 |   10019.99   673.9026    14.87   0.000     8699.164    11340.82
    _Iage_59 |   7327.841     677.23    10.82   0.000     6000.494    8655.189
    _Iage_60 |   4373.466   675.8381     6.47   0.000     3048.847    5698.086
    _Iage_61 |   1054.418   681.7956     1.55   0.122    -281.8784    2390.714
    _Iage_62 |   -3566.63   679.3208    -5.25   0.000    -4898.076   -2235.184
    _Iage_63 |  -8214.471   679.0888   -12.10   0.000    -9545.461    -6883.48
    _Iage_64 |  -11457.05   703.7792   -16.28   0.000    -12836.43   -10077.67
    _Iage_65 |  -15814.88   707.5032   -22.35   0.000    -17201.56    -14428.2
        male |   24399.33   310.8397    78.49   0.000     23790.09    25008.56
       y2010 |   31940.05   238.6503   133.84   0.000      31472.3    32407.79
  y2010_male |   12240.86   336.7094    36.35   0.000     11580.92    12900.79
       _cons |  -8687.736   592.8447   -14.65   0.000    -9849.691    -7525.78
------------------------------------------------------------------------------

. eststo c4;

. *** output results ***;
. estout *
```

```
>   using
"/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/${date}_q1fg_`sample'.txt",
>   replace
>   varwidth(12) varlabels(_cons Constant)
>   cells(b(star fmt(%12.1f)) se(par fmt(%12.1f)))
>   stats(r2 N, fmt(%9.3f %9.0g %9.3f) labels("R-squared" "Obs"));
(note: file
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/20231006_q1fg_big.txt not
found)
(output written to
/groups/brooksgrp/pppa6022/2023/problem_sets/ps2/output/20231006_q1fg_big.txt)

. ********************************************************************************
>
> ****** 2. ipums-cps data ***************************************************
>
>
********************************************************************************;
. ****** A. prep stuff ********************************************************;
. clear all;

. * set todays date;
. adopath ++ /home/lfbrooks/home/bleah/ado;
  [1]              "/home/lfbrooks/home/bleah/ado"
  [2]  (BASE)      "/usr/local/stata16/ado/base/"
  [3]  (SITE)      "/usr/local/ado/"
  [4]              "."
  [5]  (PERSONAL)  "~/ado/personal/"
  [6]  (PLUS)      "~/ado/plus/"
  [7]  (OLDPLACE)  "~/ado/"

. dateo;

. ****** B. load data ********************************************************;
. * switch for which sample we use *;
. local sample big;

. *local sample small;
. * load the big sample *;
. if "`sample'" == "big"
>   {;
.   * load cps data, keeping only variables that are of interest *;
.   use
/groups/brooksgrp/containerization/current_population_survey/input_201401/20140107_cp
sinput;
.   * take a smaller random sample *;
.   gen double rand_samp = runiform();
.   keep if rand_samp > 0.9;
(7,282,099 observations deleted)
.   * save this as a junk dataset to load *;
.   save /groups/brooksgrp/junk/ipumscps, replace;
file /groups/brooksgrp/junk/ipumscps.dta saved
.   };

. if "`sample'" == "small"
>   {;
.   * bring in the smaller data *;
.   use /groups/brooksgrp/junk/ipumscps;
.   };

. ****** C. set up data *****************************************************;
. ** fix topcodes **;
. * not in universe *;
```

```
. replace incwage = . if incwage == 9999999;
(190,346 real changes made, 190,346 to missing)

. * missing *;
. replace incwage = . if incwage == 9999998;
(153 real changes made, 153 to missing)

. * top-coded -- starts in 2011.. i think at 1,000,000 *;
. replace incwage = 1000000 if incwage == 9999997;
(0 real changes made)

. * marker for employed guys *;
. gen emp = 0;

. replace emp = 1 if empstat == 10 | empstat == 12 | empstat ==13;
(368,179 real changes made)

. ** make fake treatment **;
. * make a list of all treated states *;
. local trts 26 6 4 35 27 39 51 21 54 29 28 13 19 33 25 23;

. * make a treatment variable *;
. gen treat=0;

. * fix it for the affected states *;
. foreach j in
>   `trts'
>   {;
  2.   replace treat=1 if statefip == `j';
  3.   label variable treat "1 if ever in a treated state";
  4.   };
(21,915 real changes made)
(75,162 real changes made)
(9,172 real changes made)
(9,204 real changes made)
(11,528 real changes made)
(31,123 real changes made)
(12,025 real changes made)
(9,973 real changes made)
(8,469 real changes made)
(12,143 real changes made)
(8,054 real changes made)
(13,307 real changes made)
(10,273 real changes made)
(8,345 real changes made)
(17,794 real changes made)
(8,386 real changes made)

. * make a linear trend *;
. gen trend = year - 1979;

. * make an after variable *;
. gen after=0;

. replace after=1 if year >= 2000;
(286,463 real changes made)

. * make a trend*treatment interaction *;
. gen trend_treat = 0;

. replace trend_treat = trend * treat * after;
(98,568 real changes made)
```

```
. gen trend_treat_before = 0;

. replace trend_treat_before = trend if treat == 1;
(261,280 real changes made)

. * make a treated*after variable *;
. gen treat_after = treat*after;

. label variable treat_after "1 if after the treatment in a treated state";

. * make a treated*after*male variable *;
. gen male = 0;

. replace male = 1 if sex==1;
(391,451 real changes made)

. label variable male "1 if male (from sex)";

. gen treat_after_male = treat*after*male;

. label variable treat_after_male "1 if after the treatment in a treated state and
obs is male";

. * make other interactions that we need *;
. gen male_after = male*after;

. gen male_treat = male*treat;

. ****** C. no c now *****;
. ****** D. pictures ************************************************************;
. * keep the data around for later *;
. preserve;

. * command for making graphs i like *;
. graph set eps orientation landscape;

. *** D.1. set up data for the picture ***;
. * get annual mean wages for treated and untreated states, and also the std error *;
. sort year treat;

. collapse (mean) incwage (sem) incwage_sd = incwage if emp == 1 & year < 2000,
by(year treat);

. * reshape to make it wider *;
. reshape wide incwage incwage_sd, i(year) j(treat);
(note: j = 0 1)

Data                               long   ->   wide
-----------------------------------------------------------------------------
Number of obs.                       76   ->      38
Number of variables                   4   ->       5
j variable (2 values)             treat   ->   (dropped)
xij variables:
                                incwage   ->   incwage0 incwage1
                              incwage_sd   ->   incwage_sd0 incwage_sd1
-----------------------------------------------------------------------------

.  * calculate confidence bands *;
. foreach j in
>   0 1
>   {;
  2.   gen cihi`j' = incwage`j' + 1.96*incwage_sd`j';
```

```
  3.   label variable cihi`j' "upper confidence bound for incwage when treatment =
`j'";
  4.   gen cilo`j' = incwage`j' - 1.96*incwage_sd`j';
  5.   label variable cilo`j' "lower confidence bound for incwage when treatment =
`j'";
  6.   };

. * check this *;
. list;

     +-------------------------------------------------------------------------------
--------------------------+
     | year      incwage0   incwage_sd0      incwage1   incwage_sd1      cihi0
cilo0      cihi1      cilo1 |
     |-------------------------------------------------------------------------------
--------------------------|
  1. | 1962   3488.726801   90.37106178   3599.447927   116.9333095   3665.854
3311.6   3828.637   3370.259 |
  2. | 1963   3510.650485    93.3998764   3606.122333    127.163879   3693.714
3327.587   3855.364   3356.881 |
  3. | 1964   3709.272067   87.09248885   4064.811031   138.3090466   3879.973
3538.571   4335.897   3793.725 |
  4. | 1965       4156.32   94.96580584   4158.422072   129.5612555   4342.453
3970.187   4412.362   3904.482 |
  5. | 1966   4032.444685   61.01883199   4394.391266   87.49745929   4152.042
3912.848   4565.886   4222.896 |
     |-------------------------------------------------------------------------------
--------------------------|
  6. | 1967   4494.722594   86.70202169    4612.71028    116.099795   4664.659
4324.787   4840.266   4385.155 |
  7. | 1968   4594.673589   63.83185757   5071.785937   132.5184651   4719.784
4469.563   5331.522    4812.05 |
  8. | 1969   4927.623893   66.06521668   5378.908762   132.2097958   5057.112
4798.136    5638.04   5119.777 |
  9. | 1970   5683.795243   78.00463569   5769.535654    153.258028   5836.685
5530.906   6069.921    5469.15 |
 10. | 1971   6057.928396   83.90905401   6240.394458   167.8891473    6222.39
5893.467   6569.457   5911.332 |
     |-------------------------------------------------------------------------------
--------------------------|
 11. | 1972     6181.6962   86.52724787   6269.057526   160.9389904    6351.29
6012.103   6584.498   5953.617 |
 12. | 1973   6426.358915   88.66438436   7381.617731   212.2891514   6600.141
6252.577   7797.705   6965.531 |
 13. | 1974   6947.318742   95.03802406   7640.543323   210.3934829   7133.593
6761.044   8052.915   7228.172 |
 14. | 1975   7495.510588   103.9384367   7893.524823   233.0185656    7699.23
7291.791   8350.241   7436.809 |
 15. | 1976   7985.598623   108.5141236   9076.182426   242.5263355   8198.286
7772.911   9551.534   8600.831 |
     |-------------------------------------------------------------------------------
--------------------------|
 16. | 1977   8222.816288   115.3951966   8170.973673   151.6480878   8448.991
7996.642   8468.204   7873.744 |
 17. | 1978   9146.293871   129.0950076    9338.16241   173.5485891    9399.32
8893.268   9678.317   8998.007 |
 18. | 1979   9797.485975   135.2214121   9945.065847   177.9870414   10062.52
9532.452   10293.92   9596.211 |
 19. | 1980   10920.35409   138.6957037   10862.30014   180.5577174    11192.2
10648.51   11216.19   10508.41 |
 20. | 1981   11691.20712   142.0454523   11920.53583   198.3079201   11969.62
11412.8   12309.22   11531.85 |
```

```
     |-------------------------------------------------------------------------------
--------------------------|
 21. | 1982   12721.52574   173.3944002     13451.372   247.0383624   13061.38
12381.67   13935.57   12967.18 |
 22. | 1983   13540.21997   186.2468799   14233.84269   260.5029731   13905.26
13175.18   14744.43   13723.26 |
 23. | 1984    13949.4669   195.4333028   14775.07623   268.0979365   14332.52
13566.42   15300.55    14249.6 |
 24. | 1985   15329.11462   211.1602616   15721.13659   289.1030915   15742.99
14915.24   16287.78   15154.49 |
 25. | 1986   15957.31159   221.5348388   16566.04028   310.6661418   16391.52
15523.1   17174.95   15957.13 |
     |-------------------------------------------------------------------------------
--------------------------|
 26. | 1987   16506.42711   222.6455449     17289.852     319.98596   16942.81
16070.04   17917.03   16662.68 |
 27. | 1988    17791.3334   242.8144851   18435.76561   341.2545937   18267.25
17315.42   19104.63   17766.91 |
 28. | 1989   18347.16019   245.4002345   19184.54537   377.6299325   18828.14
17866.18    19924.7   18444.39 |
 29. | 1990   19366.22116   259.5333128   19513.10081   345.8090185   19874.91
18857.54   20190.89   18835.31 |
 30. | 1991   19603.65385   254.8610618   21012.85384   382.3647445   20103.18
19104.13   21762.29   20263.42 |
     |-------------------------------------------------------------------------------
--------------------------|
 31. | 1992   20758.66104   274.6561995   21155.98534   381.7966046   21296.99
20220.33   21904.31   20407.66 |
 32. | 1993   21182.36609   287.8277337   22102.65437     398.37453   21746.51
20618.22   22883.47   21321.84 |
 33. | 1994   22218.46129   294.4226623   22224.77471   422.0755967   22795.53
21641.39   23052.04   21397.51 |
 34. | 1995   23097.83859   317.8879622   23124.63429   425.5648989    23720.9
22474.78   23958.74   22290.53 |
 35. | 1996   25181.70613   487.4327685   26107.22981   702.4264126   26137.07
24226.34   27483.99   24730.47 |
     |-------------------------------------------------------------------------------
--------------------------|
 36. | 1997   26906.30596   557.2945621   28051.26854   790.0894449    27998.6
25814.01   29599.84   26502.69 |
 37. | 1998   27545.25089   506.5027743   26539.17301   716.5244819      28538
26552.51   27943.56   25134.79 |
 38. | 1999   28843.95846   536.0743463   29646.41378   722.5388153   29894.66
27793.25   31062.59   28230.24 |
     +-------------------------------------------------------------------------------
--------------------------+

. *** D.2. make the average wage over time picture ***;
. graph twoway
>   (rarea cilo1 cihi1 year, bcolor(gs14))
>   (rarea cilo0 cihi0 year, bcolor(gs14))
>   (connected incwage0 year, msymbol(o))
>   (connected incwage1 year, msymbol(o)),
>   ytitle("annual wage")
>   legend(off)
>   xsize(11) ysize(8.5);

. graph export "output/${date}_mean_wage_wandwo_treatment_`sample'.eps", replace;
(note: file output/20231006_mean_wage_wandwo_treatment_big.eps not found)
(file output/20231006_mean_wage_wandwo_treatment_big.eps written in EPS format)

. ****** E. regressions and summary stats
**************************************************;
```

```
. * bring back the data *;
. restore;

. *** E.1. test the parallel pre-trend assumption with a regression (2(a)) ***;
. * could have a linear trend, and an interaction term for the treated guys *;
. xi: regress incwage trend trend_treat_before i.age i.race i.statefip if emp == 1 &
year < 2000;
i.age             _Iage_0-99          (naturally coded; _Iage_0 omitted)
i.race            _Irace_100-830      (naturally coded; _Irace_100 omitted)
i.statefip        _Istatefip_1-99     (naturally coded; _Istatefip_1 omitted)
note: _Iage_1 omitted because of collinearity
note: _Iage_2 omitted because of collinearity
note: _Iage_3 omitted because of collinearity
note: _Iage_4 omitted because of collinearity
note: _Iage_5 omitted because of collinearity
note: _Iage_6 omitted because of collinearity
note: _Iage_7 omitted because of collinearity
note: _Iage_8 omitted because of collinearity
note: _Iage_9 omitted because of collinearity
note: _Iage_10 omitted because of collinearity
note: _Iage_11 omitted because of collinearity
note: _Iage_12 omitted because of collinearity
note: _Iage_13 omitted because of collinearity
note: _Iage_95 omitted because of collinearity
note: _Iage_96 omitted because of collinearity
note: _Iage_97 omitted because of collinearity
note: _Iage_98 omitted because of collinearity
note: _Iage_99 omitted because of collinearity
note: _Irace_651 omitted because of collinearity
note: _Irace_652 omitted because of collinearity
note: _Irace_801 omitted because of collinearity
note: _Irace_802 omitted because of collinearity
note: _Irace_803 omitted because of collinearity
note: _Irace_804 omitted because of collinearity
note: _Irace_805 omitted because of collinearity
note: _Irace_806 omitted because of collinearity
note: _Irace_807 omitted because of collinearity
note: _Irace_808 omitted because of collinearity
note: _Irace_809 omitted because of collinearity
note: _Irace_810 omitted because of collinearity
note: _Irace_811 omitted because of collinearity
note: _Irace_812 omitted because of collinearity
note: _Irace_813 omitted because of collinearity
note: _Irace_814 omitted because of collinearity
note: _Irace_817 omitted because of collinearity
note: _Irace_818 omitted because of collinearity
note: _Irace_820 omitted because of collinearity
note: _Irace_830 omitted because of collinearity

      Source |       SS           df       MS      Number of obs   =   232,099
-------------+----------------------------------   F(161, 231937)  =    438.69
       Model |  1.7498e+13       161  1.0869e+11   Prob > F        =    0.0000
    Residual |  5.7462e+13   231,937   247749149   R-squared       =    0.2334
-------------+----------------------------------   Adj R-squared   =    0.2329
       Total |  7.4961e+13   232,098   322969247   Root MSE        =     15740

------------------------------------------------------------------------------
           incwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------------+----------------------------------------------------------------
             trend |   714.5073   4.410046   162.02   0.000     705.8637    723.1509
trend_treat_before |  -14.68071   7.312928    -2.01   0.045    -29.01386   -.3475626
           _Iage_1 |          0  (omitted)
           _Iage_2 |          0  (omitted)
```

_Iage_3	0	(omitted)				
_Iage_4	0	(omitted)				
_Iage_5	0	(omitted)				
_Iage_6	0	(omitted)				
_Iage_7	0	(omitted)				
_Iage_8	0	(omitted)				
_Iage_9	0	(omitted)				
_Iage_10	0	(omitted)				
_Iage_11	0	(omitted)				
_Iage_12	0	(omitted)				
_Iage_13	0	(omitted)				
_Iage_14	-11337.7	15757.24	-0.72	0.472	-42221.49	19546.08
_Iage_15	-16179.95	15750.47	-1.03	0.304	-47050.47	14690.57
_Iage_16	-16703.34	15747.87	-1.06	0.289	-47568.76	14162.08
_Iage_17	-16296.18	15746.65	-1.03	0.301	-47159.21	14566.84
_Iage_18	-15150.91	15746.33	-0.96	0.336	-46013.31	15711.49
_Iage_19	-13792.38	15746.16	-0.88	0.381	-44654.45	17069.69
_Iage_20	-12639.78	15746.06	-0.80	0.422	-43501.65	18222.1
_Iage_21	-11600.32	15746	-0.74	0.461	-42462.08	19261.43
_Iage_22	-10972.19	15745.89	-0.70	0.486	-41833.72	19889.35
_Iage_23	-9876.35	15745.85	-0.63	0.531	-40737.81	20985.11
_Iage_24	-8747.714	15745.79	-0.56	0.579	-39609.06	22113.63
_Iage_25	-7176.511	15745.79	-0.46	0.649	-38037.85	23684.82
_Iage_26	-6717.766	15745.76	-0.43	0.670	-37579.05	24143.52
_Iage_27	-5800.692	15745.78	-0.37	0.713	-36662.01	25060.62
_Iage_28	-5542.695	15745.76	-0.35	0.725	-36403.98	25318.59
_Iage_29	-5035.413	15745.79	-0.32	0.749	-35896.76	25825.93
_Iage_30	-3999.315	15745.77	-0.25	0.800	-34860.61	26861.98
_Iage_31	-4135.761	15745.78	-0.26	0.793	-34997.08	26725.56
_Iage_32	-3713.418	15745.78	-0.24	0.814	-34574.75	27147.91
_Iage_33	-3650.662	15745.8	-0.23	0.817	-34512.02	27210.7
_Iage_34	-3272.118	15745.84	-0.21	0.835	-34133.56	27589.33
_Iage_35	-2303.004	15745.83	-0.15	0.884	-33164.43	28558.42
_Iage_36	-2267.219	15745.86	-0.14	0.886	-33128.71	28594.27
_Iage_37	-2068.893	15745.86	-0.13	0.895	-32930.38	28792.6
_Iage_38	-1899.566	15745.9	-0.12	0.904	-32761.13	28962
_Iage_39	-1587.141	15745.92	-0.10	0.920	-32448.73	29274.45
_Iage_40	-1753.07	15745.9	-0.11	0.911	-32614.63	29108.49
_Iage_41	-1276.694	15745.93	-0.08	0.935	-32138.3	29584.92
_Iage_42	-1447.38	15745.97	-0.09	0.927	-32309.07	29414.31
_Iage_43	-945.8378	15745.99	-0.06	0.952	-31807.56	29915.89
_Iage_44	-1101.257	15746.07	-0.07	0.944	-31963.15	29760.63
_Iage_45	-1176.723	15746.06	-0.07	0.940	-32038.6	29685.15
_Iage_46	-933.6323	15746.14	-0.06	0.953	-31795.66	29928.4
_Iage_47	-1324.826	15746.13	-0.08	0.933	-32186.84	29537.19
_Iage_48	-996.2869	15746.22	-0.06	0.950	-31858.48	29865.91
_Iage_49	-1201.365	15746.23	-0.08	0.939	-32063.57	29660.84
_Iage_50	-1579.251	15746.19	-0.10	0.920	-32441.37	29282.87
_Iage_51	-1596.409	15746.38	-0.10	0.919	-32458.91	29266.09
_Iage_52	-1852.463	15746.42	-0.12	0.906	-32715.04	29010.12
_Iage_53	-2075.492	15746.52	-0.13	0.895	-32938.26	28787.27
_Iage_54	-2186.33	15746.61	-0.14	0.890	-33049.27	28676.61
_Iage_55	-2323.98	15746.66	-0.15	0.883	-33187.02	28539.06
_Iage_56	-2860.913	15746.74	-0.18	0.856	-33724.13	28002.3
_Iage_57	-3068.174	15746.92	-0.19	0.846	-33931.74	27795.39
_Iage_58	-2707.885	15747.12	-0.17	0.863	-33571.83	28156.06
_Iage_59	-3290.393	15747.13	-0.21	0.834	-34154.37	27573.58
_Iage_60	-3414.797	15747.19	-0.22	0.828	-34278.89	27449.3
_Iage_61	-3475.769	15747.57	-0.22	0.825	-34340.61	27389.07
_Iage_62	-4773.502	15748.23	-0.30	0.762	-35639.62	26092.61
_Iage_63	-5894.772	15748.64	-0.37	0.708	-36761.71	24972.16
_Iage_64	-5480.576	15749.71	-0.35	0.728	-36349.6	25388.45
_Iage_65	-8131.596	15750.96	-0.52	0.606	-39003.07	22739.88

_Iage_66	-8579.361	15752.81	-0.54	0.586	-39454.47	22295.75
_Iage_67	-9685.14	15753.94	-0.61	0.539	-40562.46	21192.18
_Iage_68	-9089.998	15755.31	-0.58	0.564	-39970.01	21790.01
_Iage_69	-9428.931	15757.21	-0.60	0.550	-40312.66	21454.79
_Iage_70	-11088.38	15758.82	-0.70	0.482	-41975.26	19798.5
_Iage_71	-11418.23	15760.75	-0.72	0.469	-42308.9	19472.43
_Iage_72	-11088.35	15764.38	-0.70	0.482	-41986.12	19809.43
_Iage_73	-12621.12	15767.23	-0.80	0.423	-43524.5	18282.25
_Iage_74	-11776.77	15770.54	-0.75	0.455	-42686.62	19133.09
_Iage_75	-12518.82	15778.35	-0.79	0.428	-43443.98	18406.35
_Iage_76	-12346.24	15781.24	-0.78	0.434	-43277.06	18584.58
_Iage_77	-11264.68	15792.67	-0.71	0.476	-42217.9	19688.54
_Iage_78	-11951.17	15810.91	-0.76	0.450	-42940.15	19037.8
_Iage_79	-13814.92	15816.31	-0.87	0.382	-44814.47	17184.64
_Iage_80	-14126.81	15845.09	-0.89	0.373	-45182.78	16929.15
_Iage_81	-12820.57	15857.9	-0.81	0.419	-43901.65	18260.52
_Iage_82	-13298.24	15961.63	-0.83	0.405	-44582.63	17986.15
_Iage_83	-12971.88	15967.79	-0.81	0.417	-44268.33	18324.57
_Iage_84	-10967.2	16013.3	-0.68	0.493	-42352.86	20418.46
_Iage_85	-7054.943	16153.37	-0.44	0.662	-38715.13	24605.25
_Iage_86	7788.713	16338.69	0.48	0.634	-24234.7	39812.12
_Iage_87	-9237.787	16228.94	-0.57	0.569	-41046.1	22570.52
_Iage_88	-15400.63	16338.65	-0.94	0.346	-47423.97	16622.7
_Iage_89	-5797.579	17602.75	-0.33	0.742	-40298.51	28703.35
_Iage_90	-10011.72	16175.88	-0.62	0.536	-41716.03	21692.59
_Iage_91	-21387.89	19282.03	-1.11	0.267	-59180.17	16404.39
_Iage_92	-8605.467	19282.15	-0.45	0.655	-46397.99	29187.06
_Iage_93	-4680.338	22264.18	-0.21	0.833	-48317.56	38956.88
_Iage_94	-14250.44	22263.23	-0.64	0.522	-57885.79	29384.91
_Iage_95	0	(omitted)				
_Iage_96	0	(omitted)				
_Iage_97	0	(omitted)				
_Iage_98	0	(omitted)				
_Iage_99	0	(omitted)				
_Irace_200	-3266.402	122.7279	-26.61	0.000	-3506.945	-3025.858
_Irace_300	-3067.402	600.1221	-5.11	0.000	-4243.626	-1891.178
_Irace_650	1231.805	328.7511	3.75	0.000	587.461	1876.148
_Irace_651	0	(omitted)				
_Irace_652	0	(omitted)				
_Irace_700	-2175.854	291.2754	-7.47	0.000	-2746.747	-1604.962
_Irace_801	0	(omitted)				
_Irace_802	0	(omitted)				
_Irace_803	0	(omitted)				
_Irace_804	0	(omitted)				
_Irace_805	0	(omitted)				
_Irace_806	0	(omitted)				
_Irace_807	0	(omitted)				
_Irace_808	0	(omitted)				
_Irace_809	0	(omitted)				
_Irace_810	0	(omitted)				
_Irace_811	0	(omitted)				
_Irace_812	0	(omitted)				
_Irace_813	0	(omitted)				
_Irace_814	0	(omitted)				
_Irace_817	0	(omitted)				
_Irace_818	0	(omitted)				
_Irace_820	0	(omitted)				
_Irace_830	0	(omitted)				
_Istatefip_2	5483.777	468.8128	11.70	0.000	4564.916	6402.638
_Istatefip_4	1118.294	469.5953	2.38	0.017	197.8991	2038.688
_Istatefip_5	-2622.905	480.6507	-5.46	0.000	-3564.968	-1680.842
_Istatefip_6	2463.18	346.3796	7.11	0.000	1784.285	3142.075
_Istatefip_8	2466.822	451.4416	5.46	0.000	1582.008	3351.636

_Istatefip_9	3195.643	451.0901	7.08	0.000	2311.518	4079.768
_Istatefip_10	2786.736	507.9319	5.49	0.000	1791.203	3782.269
_Istatefip_11	6955.836	507.6279	13.70	0.000	5960.898	7950.773
_Istatefip_12	557.919	367.447	1.52	0.129	-162.2678	1278.106
_Istatefip_13	1337.721	420.5913	3.18	0.001	513.3731	2162.069
_Istatefip_15	1790.332	526.8991	3.40	0.001	757.6236	2823.041
_Istatefip_16	-2161.041	477.9832	-4.52	0.000	-3097.876	-1224.206
_Istatefip_17	3066.265	362.2808	8.46	0.000	2356.204	3776.326
_Istatefip_18	964.9299	405.1333	2.38	0.017	170.8791	1758.981
_Istatefip_19	-1212.43	450.9559	-2.69	0.007	-2096.292	-328.5678
_Istatefip_20	405.4579	461.5241	0.88	0.380	-499.1175	1310.033
_Istatefip_21	364.1726	451.3694	0.81	0.420	-520.4998	1248.845
_Istatefip_22	1496.469	445.9257	3.36	0.001	622.4658	2370.472
_Istatefip_23	-1175.429	503.2463	-2.34	0.020	-2161.778	-189.0787
_Istatefip_24	4060.917	424.0838	9.58	0.000	3229.723	4892.11
_Istatefip_25	3884.698	385.1406	10.09	0.000	3129.833	4639.564
_Istatefip_26	3674.966	384.4445	9.56	0.000	2921.464	4428.467
_Istatefip_27	714.8649	444.2293	1.61	0.108	-155.813	1585.543
_Istatefip_28	-720.198	477.7772	-1.51	0.132	-1656.629	216.2331
_Istatefip_29	685.7995	424.1572	1.62	0.106	-145.5377	1517.137
_Istatefip_30	-3240.445	474.8876	-6.82	0.000	-4171.212	-2309.678
_Istatefip_31	-1647.108	460.023	-3.58	0.000	-2548.741	-745.4746
_Istatefip_32	1739.198	482.3559	3.61	0.000	793.7925	2684.603
_Istatefip_33	1541.224	512.1212	3.01	0.003	537.4798	2544.968
_Istatefip_34	4542.584	370.6052	12.26	0.000	3816.207	5268.961
_Istatefip_35	-1287.493	460.5297	-2.80	0.005	-2190.12	-384.8671
_Istatefip_36	3038.823	350.7042	8.66	0.000	2351.451	3726.194
_Istatefip_37	541.7263	386.8356	1.40	0.161	-216.4614	1299.914
_Istatefip_38	-2788.337	469.9028	-5.93	0.000	-3709.334	-1867.34
_Istatefip_39	2340.619	364.4237	6.42	0.000	1626.358	3054.88
_Istatefip_40	-764.5244	466.631	-1.64	0.101	-1679.109	150.0602
_Istatefip_41	468.784	459.3095	1.02	0.307	-431.4507	1369.019
_Istatefip_42	1656.77	361.7872	4.58	0.000	947.6764	2365.864
_Istatefip_44	1461.49	512.2188	2.85	0.004	457.5545	2465.426
_Istatefip_45	450.3249	474.2768	0.95	0.342	-479.2453	1379.895
_Istatefip_46	-4007.142	460.618	-8.70	0.000	-4909.941	-3104.343
_Istatefip_47	59.17401	436.5233	0.14	0.892	-796.4004	914.7484
_Istatefip_48	843.5148	356.7958	2.36	0.018	144.2042	1542.825
_Istatefip_49	468.4996	466.1519	1.01	0.315	-445.1462	1382.145
_Istatefip_50	-1450.35	520.9992	-2.78	0.005	-2471.495	-429.2052
_Istatefip_51	4128.076	432.499	9.54	0.000	3280.389	4975.763
_Istatefip_53	2225.136	463.0088	4.81	0.000	1317.651	3132.621
_Istatefip_54	-538.3529	494.7836	-1.09	0.277	-1508.116	431.4101
_Istatefip_55	1578.483	433.1343	3.64	0.000	729.5509	2427.415
_Istatefip_56	-58.79065	496.5236	-0.12	0.906	-1031.964	914.3828
_Istatefip_61	5424.418	2682.866	2.02	0.043	166.0693	10682.77
_Istatefip_65	4115.968	3300.598	1.25	0.212	-2353.118	10585.05
_Istatefip_68	8203.318	4083.347	2.01	0.045	200.0641	16206.57
_Istatefip_69	2842.906	2682.972	1.06	0.289	-2415.65	8101.462
_Istatefip_70	2082.646	513.203	4.06	0.000	1076.781	3088.51
_Istatefip_71	2262.021	427.1177	5.30	0.000	1424.882	3099.161
_Istatefip_72	1487.486	610.2507	2.44	0.015	291.4102	2683.561
_Istatefip_73	787.8619	666.6869	1.18	0.237	-518.8272	2094.551
_Istatefip_74	2474.2	695.1516	3.56	0.000	1111.721	3836.679
_Istatefip_75	1315.839	562.74	2.34	0.019	212.883	2418.795
_Istatefip_76	697.4945	539.9557	1.29	0.196	-360.8049	1755.794
_Istatefip_77	806.0286	772.877	1.04	0.297	-708.7905	2320.848
_Istatefip_78	1632.107	703.9747	2.32	0.020	252.335	3011.88
_Istatefip_79	1635.061	773.0684	2.12	0.034	119.8664	3150.255
_Istatefip_80	2685.034	727.7463	3.69	0.000	1258.67	4111.398
_Istatefip_81	-563.8487	847.7687	-0.67	0.506	-2225.454	1097.756
_Istatefip_83	616.4479	643.3263	0.96	0.338	-644.455	1877.351
_Istatefip_84	81.41925	660.1122	0.12	0.902	-1212.384	1375.222

```
     _Istatefip_85 |  -337.2696   631.9493    -0.53   0.594    -1575.874    901.3347
     _Istatefip_87 |   .6503894   498.4765     0.00   0.999    -976.3508    977.6516
     _Istatefip_88 |   1020.867   676.0488     1.51   0.131    -304.1716    2345.905
     _Istatefip_89 |   258.3945   603.4876     0.43   0.669    -924.4257    1441.215
     _Istatefip_90 |   1767.638   558.7687     3.16   0.002     672.4654     2862.81
     _Istatefip_99 |   5635.887    3950.08     1.43   0.154    -2106.167    13377.94
             _cons |   15286.39   15747.89     0.97   0.332    -15579.07    46151.85
------------------------------------------------------------------------------

. eststo t1;

. * could test whether year effects are equal before *;
. xi: regress incwage i.year*treat i.race i.age i.statefip if emp == 1 & year < 2000;
i.year            _Iyear_1962-2013    (naturally coded; _Iyear_1962 omitted)
i.year*treat      _IyeaXtre_#         (coded as above)
i.race            _Irace_100-830      (naturally coded; _Irace_100 omitted)
i.age             _Iage_0-99          (naturally coded; _Iage_0 omitted)
i.statefip        _Istatefip_1-99     (naturally coded; _Istatefip_1 omitted)
note: _Iyear_2000 omitted because of collinearity
note: _Iyear_2001 omitted because of collinearity
note: _Iyear_2002 omitted because of collinearity
note: _Iyear_2003 omitted because of collinearity
note: _Iyear_2004 omitted because of collinearity
note: _Iyear_2005 omitted because of collinearity
note: _Iyear_2006 omitted because of collinearity
note: _Iyear_2007 omitted because of collinearity
note: _Iyear_2008 omitted because of collinearity
note: _Iyear_2009 omitted because of collinearity
note: _Iyear_2010 omitted because of collinearity
note: _Iyear_2011 omitted because of collinearity
note: _Iyear_2012 omitted because of collinearity
note: _Iyear_2013 omitted because of collinearity
note: _IyeaXtre_2000 omitted because of collinearity
note: _IyeaXtre_2001 omitted because of collinearity
note: _IyeaXtre_2002 omitted because of collinearity
note: _IyeaXtre_2003 omitted because of collinearity
note: _IyeaXtre_2004 omitted because of collinearity
note: _IyeaXtre_2005 omitted because of collinearity
note: _IyeaXtre_2006 omitted because of collinearity
note: _IyeaXtre_2007 omitted because of collinearity
note: _IyeaXtre_2008 omitted because of collinearity
note: _IyeaXtre_2009 omitted because of collinearity
note: _IyeaXtre_2010 omitted because of collinearity
note: _IyeaXtre_2011 omitted because of collinearity
note: _IyeaXtre_2012 omitted because of collinearity
note: _IyeaXtre_2013 omitted because of collinearity
note: _Irace_651 omitted because of collinearity
note: _Irace_652 omitted because of collinearity
note: _Irace_801 omitted because of collinearity
note: _Irace_802 omitted because of collinearity
note: _Irace_803 omitted because of collinearity
note: _Irace_804 omitted because of collinearity
note: _Irace_805 omitted because of collinearity
note: _Irace_806 omitted because of collinearity
note: _Irace_807 omitted because of collinearity
note: _Irace_808 omitted because of collinearity
note: _Irace_809 omitted because of collinearity
note: _Irace_810 omitted because of collinearity
note: _Irace_811 omitted because of collinearity
note: _Irace_812 omitted because of collinearity
note: _Irace_813 omitted because of collinearity
note: _Irace_814 omitted because of collinearity
note: _Irace_817 omitted because of collinearity
```

```
note: _Irace_818 omitted because of collinearity
note: _Irace_820 omitted because of collinearity
note: _Irace_830 omitted because of collinearity
note: _Iage_1 omitted because of collinearity
note: _Iage_2 omitted because of collinearity
note: _Iage_3 omitted because of collinearity
note: _Iage_4 omitted because of collinearity
note: _Iage_5 omitted because of collinearity
note: _Iage_6 omitted because of collinearity
note: _Iage_7 omitted because of collinearity
note: _Iage_8 omitted because of collinearity
note: _Iage_9 omitted because of collinearity
note: _Iage_10 omitted because of collinearity
note: _Iage_11 omitted because of collinearity
note: _Iage_12 omitted because of collinearity
note: _Iage_13 omitted because of collinearity
note: _Iage_95 omitted because of collinearity
note: _Iage_96 omitted because of collinearity
note: _Iage_97 omitted because of collinearity
note: _Iage_98 omitted because of collinearity
note: _Iage_99 omitted because of collinearity
note: _Istatefip_23 omitted because of collinearity

      Source |       SS           df       MS      Number of obs   =   232,099
-------------+----------------------------------   F(233, 231865)  =    309.74
       Model |  1.7793e+13       233  7.6366e+10   Prob > F        =    0.0000
    Residual |  5.7167e+13   231,865   246553522   R-squared       =    0.2374
-------------+----------------------------------   Adj R-squared   =    0.2366
       Total |  7.4961e+13   232,098   322969247   Root MSE        =     15702

------------------------------------------------------------------------------
       incwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------+--------------------------------------------------------------
   _Iyear_1963 |   177.0417   608.2739     0.29   0.771    -1015.159    1369.243
   _Iyear_1964 |   462.1032   538.2695     0.86   0.391    -592.8912    1517.098
   _Iyear_1965 |   848.0491   537.4168     1.58   0.115    -205.2739    1901.372
   _Iyear_1966 |    1008.97   467.6612     2.16   0.031      92.3657    1925.574
   _Iyear_1967 |   1489.552   504.9796     2.95   0.003     499.8045    2479.299
   _Iyear_1968 |    1401.41   461.9765     3.03   0.002     495.9484    2306.872
   _Iyear_1969 |   1771.505   461.9598     3.83   0.000     866.0762    2676.935
   _Iyear_1970 |   2375.963   464.3709     5.12   0.000     1465.808    3286.118
   _Iyear_1971 |   2901.807   464.0543     6.25   0.000     1992.273    3811.342
   _Iyear_1972 |   3199.713   464.8544     6.88   0.000      2288.61    4110.815
   _Iyear_1973 |    3698.96   466.0688     7.94   0.000     2785.477    4612.443
   _Iyear_1974 |   4154.678   468.5899     8.87   0.000     3236.253    5073.102
   _Iyear_1975 |   4824.055   472.2205    10.22   0.000     3898.515    5749.595
   _Iyear_1976 |   5147.984   469.3988    10.97   0.000     4227.974    6067.993
   _Iyear_1977 |   5983.987    459.135    13.03   0.000     5084.094     6883.88
   _Iyear_1978 |   6974.223   458.1331    15.22   0.000     6076.294    7872.152
   _Iyear_1979 |    7564.36   459.9886    16.44   0.000     6662.794    8465.925
   _Iyear_1980 |   8915.469   448.5875    19.87   0.000     8036.249    9794.689
   _Iyear_1981 |    9515.41   447.9503    21.24   0.000     8637.439    10393.38
   _Iyear_1982 |   10430.61   454.7093    22.94   0.000     9539.387    11321.82
   _Iyear_1983 |   10949.68   456.3117    24.00   0.000     10055.32    11844.03
   _Iyear_1984 |    11451.3   454.9087    25.17   0.000     10559.69    12342.91
   _Iyear_1985 |   12764.52   451.9587    28.24   0.000     11878.69    13650.35
   _Iyear_1986 |   13280.92   453.2082    29.30   0.000     12392.65     14169.2
   _Iyear_1987 |   13706.18   452.8559    30.27   0.000      12818.6    14593.77
   _Iyear_1988 |    14977.3   453.1025    33.05   0.000     14089.23    15865.37
   _Iyear_1989 |   15398.31   454.0158    33.92   0.000     14508.45    16288.17
   _Iyear_1990 |   16429.81   451.6533    36.38   0.000     15544.59    17315.04
   _Iyear_1991 |   16554.52   452.8372    36.56   0.000     15666.97    17442.07
   _Iyear_1992 |   17635.53   454.0466    38.84   0.000     16745.61    18525.45
```

_Iyear_1993	18017.89	453.449	39.74	0.000	17129.15	18906.64
_Iyear_1994	19047.43	453.7476	41.98	0.000	18158.1	19936.76
_Iyear_1995	20037.1	455.0305	44.03	0.000	19145.25	20928.95
_Iyear_1996	22165.24	462.8619	47.89	0.000	21258.05	23072.44
_Iyear_1997	23761.1	461.5773	51.48	0.000	22856.42	24665.78
_Iyear_1998	24402.23	459.6885	53.08	0.000	23501.25	25303.21
_Iyear_1999	25675.4	460.8205	55.72	0.000	24772.2	26578.6
_Iyear_2000	0	(omitted)				
_Iyear_2001	0	(omitted)				
_Iyear_2002	0	(omitted)				
_Iyear_2003	0	(omitted)				
_Iyear_2004	0	(omitted)				
_Iyear_2005	0	(omitted)				
_Iyear_2006	0	(omitted)				
_Iyear_2007	0	(omitted)				
_Iyear_2008	0	(omitted)				
_Iyear_2009	0	(omitted)				
_Iyear_2010	0	(omitted)				
_Iyear_2011	0	(omitted)				
_Iyear_2012	0	(omitted)				
_Iyear_2013	0	(omitted)				
treat	-880.454	808.311	-1.09	0.276	-2464.723	703.8148
_IyeaXtre_1963	105.8274	985.1483	0.11	0.914	-1825.038	2036.693
_IyeaXtre_1964	-167.4116	873.1668	-0.19	0.848	-1878.796	1543.973
_IyeaXtre_1965	-388.5158	881.7135	-0.44	0.659	-2116.652	1339.62
_IyeaXtre_1966	25.02609	759.7618	0.03	0.974	-1464.087	1514.14
_IyeaXtre_1967	-366.5131	825.8508	-0.44	0.657	-1985.159	1252.133
_IyeaXtre_1968	577.0695	811.5478	0.71	0.477	-1013.543	2167.682
_IyeaXtre_1969	304.9747	802.8589	0.38	0.704	-1268.608	1878.557
_IyeaXtre_1970	444.0297	823.1286	0.54	0.590	-1169.281	2057.34
_IyeaXtre_1971	250.7727	817.8786	0.31	0.759	-1352.248	1853.794
_IyeaXtre_1972	-190.5825	814.2673	-0.23	0.815	-1786.525	1405.36
_IyeaXtre_1973	-81.72434	840.5555	-0.10	0.923	-1729.191	1565.743
_IyeaXtre_1974	-43.16986	851.6817	-0.05	0.960	-1712.444	1626.104
_IyeaXtre_1975	-648.5022	852.6557	-0.76	0.447	-2319.685	1022.681
_IyeaXtre_1976	141.4042	842.7618	0.17	0.867	-1510.387	1793.196
_IyeaXtre_1977	-414.5519	750.9549	-0.55	0.581	-1886.404	1057.3
_IyeaXtre_1978	-369.6594	747.2173	-0.49	0.621	-1834.186	1094.867
_IyeaXtre_1979	-292.9568	747.8578	-0.39	0.695	-1758.739	1172.825
_IyeaXtre_1980	-607.4882	733.4175	-0.83	0.408	-2044.968	829.9913
_IyeaXtre_1981	-262.1055	734.763	-0.36	0.721	-1702.222	1178.011
_IyeaXtre_1982	64.13235	745.4111	0.09	0.931	-1396.854	1525.119
_IyeaXtre_1983	242.2087	746.026	0.32	0.745	-1219.983	1704.4
_IyeaXtre_1984	166.0992	744.4743	0.22	0.823	-1293.051	1625.25
_IyeaXtre_1985	-301.7912	739.4793	-0.41	0.683	-1751.152	1147.569
_IyeaXtre_1986	84.11556	745.6006	0.11	0.910	-1377.242	1545.473
_IyeaXtre_1987	312.9271	745.0833	0.42	0.674	-1147.417	1773.271
_IyeaXtre_1988	-152.6621	741.9619	-0.21	0.837	-1606.888	1301.564
_IyeaXtre_1989	286.7091	752.0858	0.38	0.703	-1187.36	1760.778
_IyeaXtre_1990	-535.8191	741.0434	-0.72	0.470	-1988.245	916.6068
_IyeaXtre_1991	644.9126	742.8459	0.87	0.385	-811.0462	2100.871
_IyeaXtre_1992	-379.0756	744.3066	-0.51	0.611	-1837.897	1079.746
_IyeaXtre_1993	-45.95331	744.7516	-0.06	0.951	-1505.647	1413.741
_IyeaXtre_1994	-716.6554	746.6464	-0.96	0.337	-2180.063	746.7522
_IyeaXtre_1995	-1032.823	747.9011	-1.38	0.167	-2498.69	433.0441
_IyeaXtre_1996	42.90171	763.1101	0.06	0.955	-1452.774	1538.578
_IyeaXtre_1997	474.6688	759.5635	0.62	0.532	-1014.056	1963.394
_IyeaXtre_1998	-1634.575	758.9391	-2.15	0.031	-3122.076	-147.0737
_IyeaXtre_1999	-92.01522	758.3891	-0.12	0.903	-1578.438	1394.408
_IyeaXtre_2000	0	(omitted)				
_IyeaXtre_2001	0	(omitted)				
_IyeaXtre_2002	0	(omitted)				
_IyeaXtre_2003	0	(omitted)				

_IyeaXtre_2004	0	(omitted)				
_IyeaXtre_2005	0	(omitted)				
_IyeaXtre_2006	0	(omitted)				
_IyeaXtre_2007	0	(omitted)				
_IyeaXtre_2008	0	(omitted)				
_IyeaXtre_2009	0	(omitted)				
_IyeaXtre_2010	0	(omitted)				
_IyeaXtre_2011	0	(omitted)				
_IyeaXtre_2012	0	(omitted)				
_IyeaXtre_2013	0	(omitted)				
_Irace_200	-3307.925	122.4653	-27.01	0.000	-3547.953	-3067.896
_Irace_300	-3660.098	599.4602	-6.11	0.000	-4835.025	-2485.172
_Irace_650	746.5339	329.1541	2.27	0.023	101.4003	1391.667
_Irace_651	0	(omitted)				
_Irace_652	0	(omitted)				
_Irace_700	-1788.52	291.4992	-6.14	0.000	-2359.851	-1217.189
_Irace_801	0	(omitted)				
_Irace_802	0	(omitted)				
_Irace_803	0	(omitted)				
_Irace_804	0	(omitted)				
_Irace_805	0	(omitted)				
_Irace_806	0	(omitted)				
_Irace_807	0	(omitted)				
_Irace_808	0	(omitted)				
_Irace_809	0	(omitted)				
_Irace_810	0	(omitted)				
_Irace_811	0	(omitted)				
_Irace_812	0	(omitted)				
_Irace_813	0	(omitted)				
_Irace_814	0	(omitted)				
_Irace_817	0	(omitted)				
_Irace_818	0	(omitted)				
_Irace_820	0	(omitted)				
_Irace_830	0	(omitted)				
_Iage_1	0	(omitted)				
_Iage_2	0	(omitted)				
_Iage_3	0	(omitted)				
_Iage_4	0	(omitted)				
_Iage_5	0	(omitted)				
_Iage_6	0	(omitted)				
_Iage_7	0	(omitted)				
_Iage_8	0	(omitted)				
_Iage_9	0	(omitted)				
_Iage_10	0	(omitted)				
_Iage_11	0	(omitted)				
_Iage_12	0	(omitted)				
_Iage_13	0	(omitted)				
_Iage_14	-11094.58	15720.39	-0.71	0.480	-41906.14	19716.97
_Iage_15	-15807.06	15713.59	-1.01	0.314	-46605.29	14991.18
_Iage_16	-16345.7	15711	-1.04	0.298	-47138.87	14447.46
_Iage_17	-15925.8	15709.81	-1.01	0.311	-46716.61	14865.02
_Iage_18	-14799.31	15709.49	-0.94	0.346	-45589.5	15990.88
_Iage_19	-13417.4	15709.32	-0.85	0.393	-44207.26	17372.46
_Iage_20	-12241.34	15709.2	-0.78	0.436	-43030.97	18548.3
_Iage_21	-11211.46	15709.16	-0.71	0.475	-42001.01	19578.08
_Iage_22	-10551.92	15709.04	-0.67	0.502	-41341.24	20237.4
_Iage_23	-9451.555	15709	-0.60	0.547	-40240.8	21337.69
_Iage_24	-8312.728	15708.94	-0.53	0.597	-39101.84	22476.39
_Iage_25	-6743.519	15708.94	-0.43	0.668	-37532.63	24045.6
_Iage_26	-6321.689	15708.91	-0.40	0.687	-37110.74	24467.36
_Iage_27	-5381.571	15708.95	-0.34	0.732	-36170.7	25407.56
_Iage_28	-5132.313	15708.91	-0.33	0.744	-35921.38	25656.75
_Iage_29	-4664.752	15708.94	-0.30	0.767	-35453.86	26124.36

_Iage_30	-3629.08	15708.91	-0.23	0.817	-34418.15	27159.99
_Iage_31	-3789.812	15708.94	-0.24	0.809	-34578.92	26999.3
_Iage_32	-3375.042	15708.94	-0.21	0.830	-34164.16	27414.08
_Iage_33	-3336.468	15708.95	-0.21	0.832	-34125.6	27452.67
_Iage_34	-2941.552	15709	-0.19	0.851	-33730.79	27847.68
_Iage_35	-2017.472	15708.99	-0.13	0.898	-32806.69	28771.74
_Iage_36	-2021.571	15709.02	-0.13	0.898	-32810.85	28767.71
_Iage_37	-1831.282	15709.02	-0.12	0.907	-32620.56	28957.99
_Iage_38	-1636.736	15709.04	-0.10	0.917	-32426.06	29152.59
_Iage_39	-1355.93	15709.07	-0.09	0.931	-32145.3	29433.44
_Iage_40	-1531.372	15709.05	-0.10	0.922	-32320.71	29257.97
_Iage_41	-1097.268	15709.08	-0.07	0.944	-31886.66	29692.13
_Iage_42	-1210.384	15709.11	-0.08	0.939	-31999.83	29579.07
_Iage_43	-765.875	15709.14	-0.05	0.961	-31555.39	30023.64
_Iage_44	-915.3278	15709.22	-0.06	0.954	-31705	29874.34
_Iage_45	-986.5059	15709.2	-0.06	0.950	-31776.13	29803.12
_Iage_46	-754.346	15709.3	-0.05	0.962	-31544.17	30035.47
_Iage_47	-1143.461	15709.27	-0.07	0.942	-31933.23	29646.31
_Iage_48	-788.4948	15709.38	-0.05	0.960	-31578.46	30001.48
_Iage_49	-1022.439	15709.38	-0.07	0.948	-31812.42	29767.55
_Iage_50	-1364.311	15709.34	-0.09	0.931	-32154.21	29425.59
_Iage_51	-1368.19	15709.53	-0.09	0.931	-32158.47	29422.09
_Iage_52	-1620.535	15709.59	-0.10	0.918	-32410.92	29169.85
_Iage_53	-1826.224	15709.68	-0.12	0.907	-32616.78	28964.34
_Iage_54	-1922.292	15709.76	-0.12	0.903	-32713.03	28868.44
_Iage_55	-2053.231	15709.81	-0.13	0.896	-32844.06	28737.6
_Iage_56	-2569.754	15709.89	-0.16	0.870	-33360.74	28221.23
_Iage_57	-2742.388	15710.07	-0.17	0.861	-33533.73	28048.95
_Iage_58	-2415.628	15710.27	-0.15	0.878	-33207.36	28376.11
_Iage_59	-2960.331	15710.28	-0.19	0.851	-33752.08	27831.42
_Iage_60	-3102.708	15710.34	-0.20	0.843	-33894.58	27689.16
_Iage_61	-3158.4	15710.72	-0.20	0.841	-33951.01	27634.21
_Iage_62	-4466.99	15711.36	-0.28	0.776	-35260.85	26326.87
_Iage_63	-5614.104	15711.79	-0.36	0.721	-36408.81	25180.61
_Iage_64	-5277.543	15712.85	-0.34	0.737	-36074.33	25519.24
_Iage_65	-7902.77	15714.1	-0.50	0.615	-38701.99	22896.45
_Iage_66	-8382.615	15715.94	-0.53	0.594	-39185.46	22420.23
_Iage_67	-9498.906	15717.05	-0.60	0.546	-40303.92	21306.11
_Iage_68	-8914.197	15718.44	-0.57	0.571	-39721.94	21893.54
_Iage_69	-9190.186	15720.31	-0.58	0.559	-40001.59	21621.22
_Iage_70	-10788.27	15721.98	-0.69	0.493	-41602.95	20026.41
_Iage_71	-11237.3	15723.88	-0.71	0.475	-42055.71	19581.11
_Iage_72	-10928.55	15727.5	-0.69	0.487	-41754.04	19896.94
_Iage_73	-12435.37	15730.32	-0.79	0.429	-43266.4	18395.66
_Iage_74	-11560.64	15733.65	-0.73	0.462	-42398.2	19276.91
_Iage_75	-12272.08	15741.38	-0.78	0.436	-43124.79	18580.62
_Iage_76	-12253.61	15744.35	-0.78	0.436	-43112.13	18604.92
_Iage_77	-11119.69	15755.72	-0.71	0.480	-42000.5	19761.12
_Iage_78	-11810.95	15773.92	-0.75	0.454	-42727.43	19105.54
_Iage_79	-13648.95	15779.26	-0.86	0.387	-44575.89	17278
_Iage_80	-13928.46	15808.12	-0.88	0.378	-44911.98	17055.05
_Iage_81	-12431.22	15820.81	-0.79	0.432	-43439.6	18577.16
_Iage_82	-12972.71	15924.38	-0.81	0.415	-44184.08	18238.65
_Iage_83	-12895.85	15930.42	-0.81	0.418	-44119.06	18327.36
_Iage_84	-10884.34	15975.88	-0.68	0.496	-42196.66	20427.98
_Iage_85	-7025.804	16115.61	-0.44	0.663	-38611.99	24560.38
_Iage_86	7667.539	16300.32	0.47	0.638	-24280.68	39615.76
_Iage_87	-9246.581	16191.06	-0.57	0.568	-40980.65	22487.48
_Iage_88	-15372.34	16300.53	-0.94	0.346	-47320.95	16576.27
_Iage_89	-6040.026	17561.14	-0.34	0.731	-40459.41	28379.36
_Iage_90	-9766.821	16138.12	-0.61	0.545	-41397.11	21863.47
_Iage_91	-20637.92	19237.38	-1.07	0.283	-58342.69	17066.85
_Iage_92	-9102.864	19238.11	-0.47	0.636	-46809.06	28603.33

_Iage_93	-3829.559	22212.29	-0.17	0.863	-47365.08	39705.96
_Iage_94	-12947.7	22215.77	-0.58	0.560	-56490.03	30594.63
_Iage_95	0	(omitted)				
_Iage_96	0	(omitted)				
_Iage_97	0	(omitted)				
_Iage_98	0	(omitted)				
_Iage_99	0	(omitted)				
_Istatefip_2	5786.421	467.9637	12.37	0.000	4869.224	6703.617
_Istatefip_4	2045.15	502.9676	4.07	0.000	1059.347	3030.954
_Istatefip_5	-2519.543	479.5946	-5.25	0.000	-3459.536	-1579.55
_Istatefip_6	3554.148	395.6418	8.98	0.000	2778.701	4329.596
_Istatefip_8	2558.427	450.535	5.68	0.000	1675.39	3441.464
_Istatefip_9	3261.708	450.6377	7.24	0.000	2378.47	4144.946
_Istatefip_10	2901.372	506.8916	5.72	0.000	1907.877	3894.866
_Istatefip_11	7167.077	506.8997	14.14	0.000	6173.566	8160.587
_Istatefip_12	693.4871	367.6524	1.89	0.059	-27.10215	1414.076
_Istatefip_13	2247.664	462.7164	4.86	0.000	1340.751	3154.576
_Istatefip_15	2042.229	525.9003	3.88	0.000	1011.478	3072.98
_Istatefip_16	-2019.717	477.0017	-4.23	0.000	-2954.628	-1084.806
_Istatefip_17	3148.822	362.8208	8.68	0.000	2437.703	3859.941
_Istatefip_18	1151.655	405.5383	2.84	0.005	356.8101	1946.499
_Istatefip_19	-163.3445	486.4944	-0.34	0.737	-1116.861	790.172
_Istatefip_20	473.4274	460.5217	1.03	0.304	-429.1833	1376.038
_Istatefip_21	1328.463	489.1037	2.72	0.007	369.8324	2287.094
_Istatefip_22	1443.522	445.5828	3.24	0.001	570.1913	2316.853
_Istatefip_23	0	(omitted)				
_Istatefip_24	4095.132	423.7711	9.66	0.000	3264.551	4925.712
_Istatefip_25	5041.597	427.6136	11.79	0.000	4203.485	5879.709
_Istatefip_26	4614.329	426.2712	10.82	0.000	3778.849	5449.81
_Istatefip_27	1714.255	480.6632	3.57	0.000	772.1674	2656.342
_Istatefip_28	342.4564	512.378	0.67	0.504	-661.7913	1346.704
_Istatefip_29	1616.014	465.5243	3.47	0.001	703.5985	2528.43
_Istatefip_30	-3015.391	473.9344	-6.36	0.000	-3944.29	-2086.492
_Istatefip_31	-1576.174	459.0382	-3.43	0.001	-2475.877	-676.4705
_Istatefip_32	1954.872	481.5035	4.06	0.000	1011.137	2898.606
_Istatefip_33	2671.775	542.2743	4.93	0.000	1608.931	3734.618
_Istatefip_34	4677.844	370.8029	12.62	0.000	3951.08	5404.608
_Istatefip_35	-163.4812	494.0537	-0.33	0.741	-1131.814	804.8513
_Istatefip_36	3116.61	351.1251	8.88	0.000	2428.414	3804.806
_Istatefip_37	733.1424	386.7055	1.90	0.058	-24.79037	1491.075
_Istatefip_38	-2568.521	468.9933	-5.48	0.000	-3487.736	-1649.306
_Istatefip_39	3437.174	412.4883	8.33	0.000	2628.708	4245.641
_Istatefip_40	-772.3464	465.6659	-1.66	0.097	-1685.04	140.3469
_Istatefip_41	506.6221	458.6344	1.10	0.269	-392.2896	1405.534
_Istatefip_42	1723.515	362.3906	4.76	0.000	1013.238	2433.791
_Istatefip_44	1501.711	511.1611	2.94	0.003	499.8484	2503.574
_Istatefip_45	491.1702	473.2806	1.04	0.299	-436.4477	1418.788
_Istatefip_46	-3789.807	459.7206	-8.24	0.000	-4690.848	-2888.766
_Istatefip_47	-17.52796	436.1983	-0.04	0.968	-872.4653	837.4094
_Istatefip_48	996.8795	357.0729	2.79	0.005	297.0258	1696.733
_Istatefip_49	568.9749	465.226	1.22	0.221	-342.8561	1480.806
_Istatefip_50	-1257.241	520.0238	-2.42	0.016	-2276.475	-238.0079
_Istatefip_51	5159.199	470.5361	10.96	0.000	4236.96	6081.437
_Istatefip_53	2173.001	461.969	4.70	0.000	1267.554	3078.449
_Istatefip_54	471.477	527.5421	0.89	0.371	-562.492	1505.446
_Istatefip_55	1412.29	432.1917	3.27	0.001	565.2054	2259.375
_Istatefip_56	77.5301	495.576	0.16	0.876	-893.7861	1048.846
_Istatefip_61	2102.4	2702.317	0.78	0.437	-3194.071	7398.871
_Istatefip_65	784.7765	3313.739	0.24	0.813	-5710.066	7279.619
_Istatefip_68	4591.624	4090.803	1.12	0.262	-3426.246	12609.49
_Istatefip_69	-453.5025	2702.41	-0.17	0.867	-5750.157	4843.152
_Istatefip_70	2131.326	529.5287	4.02	0.000	1093.463	3169.188
_Istatefip_71	2777.717	440.6413	6.30	0.000	1914.071	3641.363

```
 _Istatefip_72 |   1555.659   623.6629     2.49   0.013     333.2955    2778.022
 _Istatefip_73 |   824.2032   678.7307     1.21   0.225    -506.0915    2154.498
 _Istatefip_74 |    2560.37   706.6364     3.62   0.000     1175.381    3945.359
 _Istatefip_75 |   1386.993   577.5329     2.40   0.016     255.0434    2518.943
 _Istatefip_76 |   1259.795   550.3796     2.29   0.022     181.0651    2338.525
 _Istatefip_77 |   879.5233   782.8365     1.12   0.261    -654.8161    2413.863
 _Istatefip_78 |   1728.804   715.4468     2.42   0.016      326.547    3131.061
 _Istatefip_79 |   1716.581   783.1613     2.19   0.028     181.6047    3251.557
 _Istatefip_80 |   2700.578   738.7596     3.66   0.000     1252.628    4148.528
 _Istatefip_81 |   569.4249   862.0682     0.66   0.509    -1120.207    2259.056
 _Istatefip_83 |   1764.302   663.2474     2.66   0.008     464.3538    3064.249
 _Istatefip_84 |   1223.089   679.4825     1.80   0.072    -108.6796    2554.857
 _Istatefip_85 |   830.9829   652.3132     1.27   0.203    -447.5342      2109.5
 _Istatefip_87 |   1138.546   524.5986     2.17   0.030     110.3468    2166.746
 _Istatefip_88 |    2143.64   694.8944     3.08   0.002     781.6652    3505.615
 _Istatefip_89 |   1419.058   624.9288     2.27   0.023     194.2136    2643.902
 _Istatefip_90 |   2912.073   582.0368     5.00   0.000     1771.296     4052.85
 _Istatefip_99 |   3858.869   3946.374     0.98   0.328    -3875.922    11593.66
         _cons |   6162.495   15715.78     0.39   0.695    -24640.02    36965.01
------------------------------------------------------------------------------

. local testvals _IyeaXtre_1963;

. forvalues y=1964/1999
>   {;
  2.   local testvals `testvals' = _IyeaXtre_`y';
  3.   };

. test `testvals' = 0;

 ( 1)  _IyeaXtre_1963 - _IyeaXtre_1964 = 0
 ( 2)  _IyeaXtre_1963 - _IyeaXtre_1965 = 0
 ( 3)  _IyeaXtre_1963 - _IyeaXtre_1966 = 0
 ( 4)  _IyeaXtre_1963 - _IyeaXtre_1967 = 0
 ( 5)  _IyeaXtre_1963 - _IyeaXtre_1968 = 0
 ( 6)  _IyeaXtre_1963 - _IyeaXtre_1969 = 0
 ( 7)  _IyeaXtre_1963 - _IyeaXtre_1970 = 0
 ( 8)  _IyeaXtre_1963 - _IyeaXtre_1971 = 0
 ( 9)  _IyeaXtre_1963 - _IyeaXtre_1972 = 0
 (10)  _IyeaXtre_1963 - _IyeaXtre_1973 = 0
 (11)  _IyeaXtre_1963 - _IyeaXtre_1974 = 0
 (12)  _IyeaXtre_1963 - _IyeaXtre_1975 = 0
 (13)  _IyeaXtre_1963 - _IyeaXtre_1976 = 0
 (14)  _IyeaXtre_1963 - _IyeaXtre_1977 = 0
 (15)  _IyeaXtre_1963 - _IyeaXtre_1978 = 0
 (16)  _IyeaXtre_1963 - _IyeaXtre_1979 = 0
 (17)  _IyeaXtre_1963 - _IyeaXtre_1980 = 0
 (18)  _IyeaXtre_1963 - _IyeaXtre_1981 = 0
 (19)  _IyeaXtre_1963 - _IyeaXtre_1982 = 0
 (20)  _IyeaXtre_1963 - _IyeaXtre_1983 = 0
 (21)  _IyeaXtre_1963 - _IyeaXtre_1984 = 0
 (22)  _IyeaXtre_1963 - _IyeaXtre_1985 = 0
 (23)  _IyeaXtre_1963 - _IyeaXtre_1986 = 0
 (24)  _IyeaXtre_1963 - _IyeaXtre_1987 = 0
 (25)  _IyeaXtre_1963 - _IyeaXtre_1988 = 0
 (26)  _IyeaXtre_1963 - _IyeaXtre_1989 = 0
 (27)  _IyeaXtre_1963 - _IyeaXtre_1990 = 0
 (28)  _IyeaXtre_1963 - _IyeaXtre_1991 = 0
 (29)  _IyeaXtre_1963 - _IyeaXtre_1992 = 0
 (30)  _IyeaXtre_1963 - _IyeaXtre_1993 = 0
 (31)  _IyeaXtre_1963 - _IyeaXtre_1994 = 0
 (32)  _IyeaXtre_1963 - _IyeaXtre_1995 = 0
 (33)  _IyeaXtre_1963 - _IyeaXtre_1996 = 0
```

```
 (34)  _IyeaXtre_1963 - _IyeaXtre_1997 = 0
 (35)  _IyeaXtre_1963 - _IyeaXtre_1998 = 0
 (36)  _IyeaXtre_1963 - _IyeaXtre_1999 = 0
 (37)  _IyeaXtre_1963 = 0

       F( 37,231865) =    1.12
            Prob > F =    0.2776

. eststo t2;

. *** E.2. diff-in-diff regression *****;
. * regression to paralell diff in diff summary stats *;
. regress incwage treat_after treat after if emp == 1;

      Source |       SS           df       MS      Number of obs   =   365,948
-------------+----------------------------------   F(3, 365944)    =  15011.22
       Model |  4.8351e+13         3  1.6117e+13   Prob > F        =    0.0000
    Residual |  3.9290e+14   365,944  1.0737e+09   R-squared       =    0.1096
-------------+----------------------------------   Adj R-squared   =    0.1096
       Total |  4.4125e+14   365,947  1.2058e+09   Root MSE        =     32767

------------------------------------------------------------------------------
     incwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 treat_after |  -1360.156   238.5852    -5.70   0.000    -1827.776   -892.5355
       treat |   1358.759   145.6492     9.33   0.000     1073.291    1644.227
       after |   24279.31   137.7779   176.22   0.000     24009.27    24549.35
       _cons |   13608.72   82.55693   164.84   0.000     13446.91    13770.53
------------------------------------------------------------------------------

. eststo r0;

. * make a marker for the regression sample *;
. gen r0emp = e(sample);

. *** E.3. summary tables for diff-in-diff ***;
. * save the data for later *;
. preserve;

. sort treat after;

. collapse (mean) incwage (sem) incwage_sd = incwage (count) incwage_obs=incwage
>   if r0emp == 1, by(treat after);

. * output to excel *;
. outsheet using output/${date}_prob2c_ttests_`sample'.txt, replace;
(note: file output/20231006_prob2c_ttests_big.txt not found)

.  *** E.4. output regression results ***;
. estout *
>   using "output/${date}_prob2_regs_`sample'.txt",
>   replace
>   varwidth(12) varlabels(_cons Constant)
>   cells(b(star fmt(%12.1f)) se(par fmt(%12.1f)))
>   stats(r2 N, fmt(%9.3f %9.0g %9.3f) labels("R-squared" "Observations"));
(note: file output/20231006_prob2_regs_big.txt not found)
(output written to output/20231006_prob2_regs_big.txt)

.  log close;
      name:  <unnamed>
       log:  /groups/brooksgrp/pppa6022/2023_fall/problem_sets/ps2/ps2v15.log
  log type:  text
 closed on:   6 Oct 2023, 00:04:11
```