

Problem Set# 1: Does a minimum wage increase lead to an increase in the percentage of the workforce that is formally employed?

Answer Key

In this problem set, we will review the lab discussing the differences between a hypothesis test of a difference in means (Diff in Diff) and how the same test can be conducted using a multiple regression model. The focus in this problem set is train your skills in writing about a statistical analysis and its interpretation. The datasets and example refers to the example discussed in Chapter 2 of *Quantitative Social Science: An Introduction*. Please prepare a brief document including output where pertinent for your answer.

Part I. Statistical Tests and their interpretation.

- (a) List the variables in the two data sets (minwage.csv and minwage-reshaped.dta) and use Stata's sum command to produce descriptive stats. Please summarize the difference between both data sets. (Hint: you can use the describe command in Stata).

We can use Stata's describe command. The key difference to understand is that in the first data set, we have 358 observations. In the second data set, we have 716 observations or twice the amount because the observations are stacked upon each other.

Contains data from C:\Users\Lorena Barberia\Dropbox\FLS 6183\2018 Lecture Slides\Lab\minwage.dta

```
obs:      358
vars:      8                22 Aug 2018 09:06
size:     15,394
```

variable name	storage type	display format	value label	variable label
chain	str10	%10s		
location	str9	%9s		
wagebefore	float	%9.0g		wageBefore
wageafter	float	%9.0g		wageAfter
fullbefore	float	%9.0g		fullBefore
fullafter	float	%9.0g		fullAfter
partbefore	float	%9.0g		partBefore
partafter	float	%9.0g		partAfter

```
. describe
```

Contains data from C:\Users\Lorena Barberia\Dropbox\FLS 6183\2018 Lecture Slides\Lab\minwage-resshaped.dta

obs: 716
vars: 5 22 Aug 2018 09:08
size: 17,900

variable name	storage type	display format	value label	variable label
uniqueid	str12	%12s		
time	byte	%9.0g		
wage	float	%9.0g		
NJ	float	%9.0g		
fullprop	float	%9.0g		

Sorted by: uniqueid

- (b) Construct a t-test and 95 percent confidence interval for the difference in the mean rate of full-time employment in New Jersey. What are the problems with using this measure to assess the impact of the minimum wage increase on employment percentages? Why is this exercise called the naïve ttest?

```
. * Naive Test// Compute the difference-in-means NJ
. ttest fullpropbefore=fullpropafter if minwageNJ==1
```

Paired t test

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
fullpr~e	291	.2965262	.0135098	.2304592	.2699366	.3231158
fullpr~r	291	.320401	.014714	.2510016	.2914413	.3493607
diff	291	-.0238747	.0176471	.3010377	-.0586075	.010858

```
mean(diff) = mean(fullpropbefore - fullpropafter)      t = -1.3529
Ho: mean(diff) = 0                                     degrees of freedom = 290

Ha: mean(diff) < 0      Ha: mean(diff) != 0      Ha: mean(diff) > 0
Pr(T < t) = 0.0886      Pr(|T| > |t|) = 0.1771      Pr(T > t) = 0.9114
```

On average, 32% of employees are full-time in New Jersey after the wage increase as compared to only 29.65%. The t-test tells us that the difference 2.38% (an increase in full-time employment) is not statistically significant. The test is naïve because we are not comparing NJ to a counterfactual case.

- (c) Now, let's construct a t-test and 95 percent confidence interval for the difference in the average full-time employment in Pennsylvania before and after the implementation of the minimum wage increase. (Hint: This command is in the dofile, but you can easily modify the commands in the do file to run this test). What is the approximate percentage difference in formal employment in PA?

```
. ttest fullpropbefore=fullpropafter if minwagePA==1
```

Paired t test

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
fullpr~e	67	.3099657	.0292949	.2397895	.2514764	.3684549
fullpr~r	67	.2722821	.0302054	.2472422	.211975	.3325892
diff	67	.0376836	.0419521	.3433928	-.0460765	.1214437

mean(diff) = mean(fullpropbefore - fullpropafter) t = 0.8983
Ho: mean(diff) = 0 degrees of freedom = 66

Ha: mean(diff) < 0 Ha: mean(diff) != 0 Ha: mean(diff) > 0
Pr(T < t) = 0.8138 Pr(|T| > |t|) = 0.3723 Pr(T > t) = 0.1862

On average, 30.99% of employees are full-time in PA before the wage increase as compared to 27.22% afterwards. The t-test tells us that the difference 3.77% (a decrease in full-time employment) is not statistically significant.

(d) Stata's testing procedure calculates the standard error of differences in averages under two alternative assumptions about variances. State these assumptions in words. (Hint: see the help file for the specific ttest command in Stata). Which calculation corresponds to the default for a ttest? Explain and check your answer.

(e) In class, we discussed that we can calculate:

$$DiffinDiff_{t-(t-1)} = \left\{ (\bar{Y}_t - \bar{Y}_{t-1}) \middle| T=1 \right\} - \left\{ (\bar{Y}_t - \bar{Y}_{t-1}) \middle| T=0 \right\}$$

Use the information in (b) and (c) to calculate the difference the proportion of full-time employment in Pennsylvania and New Jersey before and after the implementation of the minimum wage increase.

$$DiffinDiff_{t-(t-1)} = \left\{ (32.04 - 29.65) \middle| NJ = 1 \right\} - \left\{ (30.99 - 27.22) \middle| PA = 0 \right\} = 2.38 - (-3.77) = .0615583$$

(f) Use Stata's reg command to prove the result in question (e) above "by computer." Please discuss in detail how to interpret each parameter in the regression that was estimated:

```
. reg fullprop NJ time did
```

Source	SS	df	MS	Number of obs = 716		
Model	.163256214	3	.054418738	F(3, 712) = 0.93		
Residual	41.5022705	712	.058289706	Prob > F = 0.4240		
Total	41.6655267	715	.058273464	R-squared = 0.0039		
				Adj R-squared = -0.0003		
				Root MSE = .24143		

fullprop	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
NJ	-.0134395	.0327155	-0.41	0.681	-.0776698	.0507909
time	-.0376836	.0417132	-0.90	0.367	-.1195792	.044212
did	.0615583	.0462667	1.33	0.184	-.0292772	.1523938
_cons	.3099657	.0294957	10.51	0.000	.2520568	.3678746

(g) Now, let's examine the same test using the Stata package, `diff`. Please copy the table and review the results.

```
. diff fullprop, t(NJ) p(time)
```

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS

Number of observations in the DIFF-IN-DIFF: 716

	Before	After		
Control:	67	67	134	
Treated:	291	291	582	
	358	358		

Outcome var.	fullp~p	S. Err.	t	P> t
Before				
Control	0.310			
Treated	0.297			
Diff (T-C)	-0.013	0.033	-0.41	0.681
After				
Control	0.272			
Treated	0.320			
Diff (T-C)	0.048	0.033	1.47	0.142
Diff-in-Diff	0.062	0.046	1.33	0.184

R-square: 0.00

* Means and Standard Errors are estimated by linear regression

Inference: * p<0.01; ** p<0.05; * p<0.1

"The regression results suggest that the effect was small and the confidence interval around it was wide enough that we cannot be confident even of its direction, although the majority of the confidence interval is positive." To arrive at the same conclusion, we would have to calculate the confidence interval from the Diff in Diff output.

Note, we may be hesitant to interpret interaction models based on their p-values. We may wish to look instead at the predicted values of the outcome variable in all four groups, and the marginal effect of the time variable in both treatment groups, and then make a judgment call as to whether the differences are large enough to matter from a practical perspective.

To really understand your model and its implications, let's re-run the model using factor variable notation. See `-help fvvarlist-` for details.

Code:

```
regress fullprop i.time##i.NJ, robust
```

And then we can run `-margins-`:

Code:

```
margins time#treated
```

```
margins treated, dydx(time)
```

which will show us the expected values of `fullprop` in each combination of time and treat, and then the change in `fullprop` over time in both groups.

Now we can compare the marginal effects of time in each group.

Part II. Research Designs

- (h) Compare the results obtained using the *difference-in-difference* approach and the multiple regression model. What are the similarities and differences?

Even though DID analyses do a somewhat better job of identifying causal effects than simpler cohort designs, they are not perfect in that regard.

“Usando o método estatístico diff in diff, com a mesma interação entre tratamento local*tempo obtemos resultados semelhantes ao da regressão. O coeficiente, por exemplo, é o mesmo 0.062, bem como a mesma significância estatística.”

The R2 is very low in the regression model. Why is this the case?

<https://www.statalist.org/forums/forum/general-stata-discussion/general/1379375-need-help-on-result-of-difference-in-differences-estimator>

“Na abordagem diff in diff não temos o quanto a variável de controle poderia afetar a variável de tratamento ($X \leftarrow Y$), enquanto que na regressão multivariada conseguiríamos incluir variáveis de controle para resolver o problema da confounder bias. Outro problema que não é alcançado pelo diff in diff é o da endogeneidade, que poderia ser resolvido facilmente na regressão pela inclusão de outras variáveis. Por fim, o componente estocástico (erro - resíduo) só é mensurado via regressão.”

~

“Primeiramente, nota-se que o coeficiente relatado pelo diff-in-diff corresponde ao mesmo valor relatado pela coeficiente da interação, ou seja, 0,062. Nesse comando, também são apresentadas todas as diferenças que foram calculadas anteriormente. Tendo em vista o valor do p valor apresentado para o coeficiente do diff-in-diff, temos que ele apenas apresenta significância estatística assumindo um nível de confiança de 90%, valor distante do padrão utilizado para que podemos rejeitar a hipótese nula.”

- (i) Can you think about any confounders that might affect the results?

“Políticas macroeconômicas poderiam afetar a relação encontrada entre X e Y, bem como, a associação do salário ao gênero (dado que mulheres ganham menos) poderia afetar negativamente a média de um local com mais mulheres que homens; a existência ou não de outras políticas que incentivem a jornada de trabalho completa e.g. subsídio de creche. Entre outras.”

“A semelhança entre os dois métodos é de que o coeficiente de interação representa o mesmo que o coeficiente do diff-in-diff. A diferença é de que a regressão permite que sejam adicionadas variáveis de controle na análise. Os resultados apontaram de forma clara que o aumento do salário mínimo não produziu efeitos sobre o desemprego no estado de NJ. A redução de desemprego pode ter sido resultado de outras variáveis que não foram captadas pelo modelo de regressão. Nesse método, a solução seria a utilização de outras variáveis. Porém no caso do método de diff-in-diff não é possível controlar as demais variações que podem acontecer ao longo do tempo.”

“Variáveis confounders que podem afetar a relação entre a variável dependente e independente podem estar relacionadas a situação econômica dos estados. Além disso, por mais que tratem-se de estado vizinhos, integram um país que permite ampla autonomia no estabelecimento das leis. Nesse sentido, o impacto de uma diferente legislação trabalhista pode exercer um importante efeito sobre o resultado do desemprego.”

- (j) What are some of the limitations of *before-and-after designs* (comparison between pre- and posttreatment measurements)?

“A dificuldade da existência de contrafactual como foi possível neste exemplo. E, mais que isso, a série temporal impõe dificuldades adicionais à regressão, já que a não aleatoriedade (ou, noutras palavras, dificuldade em identidade nos grupos em T0 e T1) faz com que a distribuição dos grupos no tempo 0 e 1 não seja a mesma.”

“A principal limitação na utilização do método de diff-in-diff está na impossibilidade da utilização de variáveis que controlem a variação ao longo do tempo. O efeito da variável salário sobre o desemprego em cada um desses estados pode estar relacionada à variáveis específicas aos seus contextos que podem ou não estar presentes em ambos os estados. Na ausência da utilização variáveis que controlem demais variáveis que também podem produzir efeito sobre a dependente, a relação encontrada pelo diff-in-diff resulta em valores não significativos. Para que esse método seja utilizado é preciso que os casos de controle e tratamento sejam estritamente parecidos, ou seja, a variação deve apenas ser explicada pelas variáveis propostas pelo modelo. No caso apresentado, o diff-in-diff poderia apenas ser útil caso o desemprego fosse apenas explicado pelo salário mínimo, o que não é o caso em nenhum dos dois estados.”

Part III. Random Assignment

A professor decides to run an experiment to measure the effect of time pressure on final exam scores. He gives each of his 400 students in his course the same final exam, but some students have 90 minutes to complete the exam while others have 120 minutes. Each student is randomly assigned one of the examination times based on a flip of the coin. Let Y_i denote the number of points scored on the exam by the i th student ($0 \leq Y_i \leq 100$), let X_i denote the amount of time the student has to complete the exam ($X_i=90$ or 120), and consider the regression model $Y_i = \beta_0 + \beta_1 X_i + u_i$.

- (k) Explain what the term u_i represents. Why will different students have different values of u_i ?

u_i represents factors other than time that influence the student's performance on the exam including amount of time studying, aptitude for the material, and so forth. Some students will have studied more than average, other less; some students will have higher than average aptitude for the subject, others lower, and so forth.

- (l) Explain why $E(u_i|x_i)=0$ for this regression model. Are the other assumptions mentioned in Stock and Watson Chapter 4.3 satisfied? Explain.

Because of random assignment u_i is independent of X_i . Since u_i represents deviations from average $E(u_i) = 0$. Because u and X are independent $E(u_i|X_i) = E(u_i) = 0$.

- (m) The estimated regression is $Y_i = 49 + 0.24X_i + u_i$.
1. Compute the estimated regression's prediction for the average score of students given 90 minutes to complete the exam. Repeat for 120 minutes and 150 minutes.

$$Y_i = 49 + 0.24(90) = 49 + 0.24 * 90 = 70.6$$

$$Y_i = 49 + 0.24(120) = 77.8$$

$$Y_i = 49 + 0.24(150) = 85$$

2. Compute the estimated gain in score for a student who is given an additional 10 minutes on the exam.

$$\frac{\Delta Y_i}{\Delta X_i} = 0.24 \text{ and therefore } 0.24 * 10 = 2.4$$