



**MPPs crawling to the
end of first year**

**The last time I felt this emotionally encumbered, I was playing Lady Macbeth
on a Crystal Skies cruise ship during Shakespeare at Sea Week.**

Schitt's Creview

API 203: TF Exam Review

R

**Nolan M. Kavanagh
May 3, 2024**

Practice questions!

You work for the governor of Texas, where the minimum wage is \$7.25. He is considering raising it but needs more information.

He asks you to evaluate how minimum wage increases in other states have affected their labor markets. You have data on all states' minimum wages and labor markets from 2011 to 2022, and you run a regression with year and state fixed effects.

Which omitted variables does this approach eliminate?

- A. State demographics that don't change over time.**
- B. State demographics that do change over time.**
- C. National trends in economic growth.**
- D. National economic policies, e.g. labor laws.**
- E. Pennsylvania economic policies that were passed in 1998.**
- F. Pennsylvania economic policies that were passed in 2014.**

Practice questions!

You take a closer look at the minimum wages of each state to better understand your data.

You notice the following minimum wages (see table). You can assume that if two years are the same, the minimum wage didn't change between them.

	2011	2016	2022
State 1	\$7.25	\$7.25	\$7.25
State 2	\$8.25	\$10.35	\$10.35
State 3	\$12.75	\$12.75	\$12.75
State 4	\$7.25	\$9.75	\$10.35

Which of the following state(s) is/are contributing variation to your fixed effects regression?

- A. State 1**
- B. State 2**
- C. State 3**
- D. State 4**

Practice questions!

You run a regression of $\log(\text{labor market size})$ (with the market size measured in millions of people) on a state's minimum wage (in \$USD) with state and year fixed effects from 2011 to 2022.

You get an estimate of $\beta_1 = 0.06$ (95% CI, 0.02 to 0.10).

Which of the following is the best interpretation of this estimate?

- A. A \$1 increase in a state's minimum wage was associated with a 6 percentage point increase in the state's labor market.**
- B. A \$1 increase in a state's minimum wage was associated with a 60,000-person increase in the state's labor market.**
- C. A \$1 increase in a state's minimum wage was associated with a 0.06 percent increase in the state's labor market.**
- D. A \$1 increase in a state's minimum wage was associated with a 6% increase in the state's labor market.**
- E. A 1% increase in a state's minimum wage was associated with a 600-person increase in the state's labor market.**

Practice questions!

The governor asks you to broaden your search for any evidence of increased take-home pay on worker well-being. You notice that people at 100–400% of the federal poverty level (FPL) are eligible for health insurance subsidies, but people living <100% of FPL are not.

You believe that this context is perfect for regression discontinuity. Which of the following lines of code estimates the causal effect of eligibility for insurance subsidies on workers' financial stress?

A. `feols(stress ~ FPL + above_FPL_100 | 0, data)`

B. `feols(stress ~ FPL + above_FPL_100 + FPL*above_FPL_100 | FPL, data)`

C. `feols(stress ~ above_FPL_100 | 0, data)`

D. `feols(stress ~ FPL + above_FPL_100 + FPL*above_FPL_100 | 0, data)`

Practice questions!

You retrieve the following RD OLS estimates. Note that financial stress is measured as a binary, with 0 indicating no financial stress and 1 indicating financial stress.

Which is the most accurate interpretation of the findings?

- A. Being eligible for the insurance subsidies causes a 10.2 percentage point decrease in the probability of financial stress. The result is statistically significant at the 5% level.
- B. Being eligible for the insurance subsidies causes a 10.2 percentage point increase in the probability of financial stress. The result is statistically significant at the 5% level.
- C. Being eligible for the subsidies causes a 2.3 percentage point decrease in the probability of financial stress, but the difference is not statistically significant at the 5% level.
- D. Every 1 percentage point increase in FPL causes a 0.3 percent decline in the probability of being financially stressed. The result is not statistically significant at the 5% level.

	Model 1
Intercept	0.601 (0.023)
FPL	-0.003 (0.009)
above_FPL_100	-0.102 (0.035)
FPL * above_FPL_100	-0.023 (0.029)
Num.Obs.	2,345
R2	0.034
R2 Adj.	0.032

Practice questions!

Describe two criteria for the subsidy eligibility threshold that must be met to estimate a causal effect on financial stress. Briefly evaluate whether you believe these criteria were met here.

Practice questions!

To whom does your causal estimate apply?

Recall that you work for the governor of Texas, who is interested in raising the minimum wage.

Describe how your results may or may not generalize to your governor's constituents.

Practice questions!

The governor (surprisingly) decides to pilot a universal basic income for households earning less than the federal poverty level. The available funds are limited, so you randomize it among those who apply. However, only about 50% of lottery winners end up getting the money.

You're still interested in the effect of the universal basic income on financial stress.

What's a statistical approach you could use to causally answer this question? Explain the strengths and weaknesses of your approach.

Practice questions!

You decide to pursue an instrumental variables approach to answer this causal question.

Write out the regressions for the following three models, where the Y variable is “stress”, D is “got_money”, and Z is “won_lottery”:

1. The reduced form

2. The first stage

3. The second stage

Indicate which coefficient(s) give the causal estimate that you’re interested in.

Practice questions!

Which of the following are potential threats to the internal validity of your IV estimate?

Select all that apply.

A. The group of people who enrolled in the lottery were more white and lower-income than the people who would get the policy once it's fully rolled out.

B. There is evidence that the administrators of the program gave preference to people living in major cities when selecting the lottery winners.

C. Some people won the lottery and didn't end up getting the money, but they did enroll in other government programs they were eligible for.

D. None of the above are threats to inference.

Practice questions!

The governor asks you to identify Texans with the most financial stress so that the administration can design other programs to help them.

You don't have the resources to collect a survey of all Texans, so you decide to make a predictive model using your existing data.

You take two approaches: an OLS model with all predictor variables that you have available, and a LASSO model using the same predictors.

Which of the following statements is most likely to be true?

- A. The OLS model will probably have the greatest out-of-sample prediction because it uses all available variables.
- B. The LASSO model uses a tuning parameter to select highly correlated variables since including predictors that are highly correlated with each other tends to improve a model's predictive power.
- C. We can't be sure which model is necessarily more predictive without evaluating recall and precision, preferably in the test set.
- D. The OLS model is likely to be overfitted, but we can evaluate overfitting using the model's recall and precision in our training set.

Practice questions!

The conservative governor is worried about the political consequences of giving “handouts” to people who don’t actually need help.

As a result, he asks you to select a model based on how well it correctly identifies people in need — without including financially healthy people.

Which of the following measures is most directly useful for this purpose?

- A. Specificity**
- B. Recall (a.k.a. sensitivity)**
- C. Precision (a.k.a. positive predictive value)**
- D. Negative predictive value**

Good luck!