Assignment 4
Yein Shin

- **What implicit claim about causality does Obama's "cycle of crime" theory assert?**

The cycle of crime theory asserts the causality that incarceration leads to recidivism since criminals may find it hard to get a new job after incarceration due to background checks and get along with other criminals, getting new criminal skills.

- **Your friend has an ingenious idea. He/she has detailed case data about criminal sentencing in a large jurisdiction for everyone charged with a felony. The data includes the length of the prison sentence (in days), and whether the person was convicted of a second crime after he/she was out ("recidivism"). This seems to be what the "cycle of crime" theory is talking about.**

  The proposed research design is: ***Run a regression whose outcome is recidivism and whose main explanatory variable is the length of the prison sentence.***

  React your friend's research design.

I would react to my friend's research design by pointing out some methodological issues regarding causal inference. First, the proposed research design is vulnerable to a reverse causality (or simultaneity) issue. Now, the researcher is examining the relationship between the length of the prison sentence (Independent variable, x) and recidivism (Dependent variable, y), but it is also likely that criminals (or a certain crime) who are expected to have a higher possibility of committing crime again (e.g., who never regret their wrongdoings) get a longer sentence. Also, some criminals with good behavior get an early release from prison. That is, it is likely that when a criminal has a lower possibility of recidivism, they will spend less time in prison. If this is the case, then it is possible that not only the length of the prison sentence(x) affects recidivism (y) but also expected recidivism rates (y) associated with a certain crime or criminal explains the length of the prison sentence (x) as well.

Second, there are so many potential omitted variables that might distort the true relationship between the length of the prison sentence and recidivism. For example, a criminal's family background or economic situation affects both the prison sentence length (x) and recidivism (y). A criminal may end up getting a longer sentence since they cannot afford the lawyer's fee and may commit a crime again for money. Also, if a criminal is an illegal alien, they may get a longer sentence due to aggravated punishment, and even after the release, they may still find it hard to get a proper job that might keep them out of crime. Thus, 'less privileged' criminals will get a longer sentence (x) and recidivate (y) more.

Lastly and most importantly, this research design has a serious logical flaw. Even though the suggested mechanism (i.e., criminals who spend more time in prison are more likely to recidivate because they may learn new criminal skills from other criminals in jail) is true, the relationship can hardly be proved. If the criminal upgrades his criminal skills, he has less chance of getting caught by police as the crime becomes more sophisticated (more so since criminals who spent terrible time in jail would desperately wish to avoid going to jail again). If a statistically significant and positive relationship exists between the length of the prison sentence and recidivism, then my last critique will enhance the researcher's claim since he found a significantly positive relationship even with the above-mentioned difficulty. However, if the researcher fails to find a significantly positive relationship, then he may want to rethink my claim. The research design makes it harder for a researcher to prove the mechanism he wants to assert. If the criminals have learned new criminal skills in jail, they may recidivate more but have less chance of getting caught. Since the data is only for defendants who went to trial (who got caught), my critique may be valid.

I also got some recommendations about the research question. My recommendations would hardly mitigate the above concerns, but first, it would be helpful to claim a mechanism if there exists a variation regarding the enforcement of background checks depending on state, industry, or job. It would be better if there were some unexpected changes in the enforcement of background checks. Then, we can try difference-in-differences, which may mitigate the reverse causality issue. A good instrumental variable is also welcomed since the research design is vulnerable to omitted variable bias.

- **You will now develop a separate research design. You notice that the data contains names and identifiers for judges. By merging in with a separate dataset, you are able to add a new variable representing whether each defendant's presiding judge was appointed by a Democrat or a Republican. You also learn that judges in this jurisdiction (and in most) are randomly assigned to defendants. Perform a balance test. Does the judge's party really seem to be randomly assigned?**

When it comes to the severity of crime, the difference between control (i.e., the judge was appointed by a Democrat) and treatment (i.e., the judge was appointed by a Republican) is statistically insignificant, according to Table 1. Thus, it can be said that there's no tendency for more severe crimes to be assigned to a Democrat judge or vice versa. For the severity of crime, the judges are randomly assigned.

Table 1. Balance test

	Control	**Treatment**	**Difference**
Severity of crime	1.979	1.966	0.014

- **Describe in words the ``first stage'' of the IV design. Then, create a publication-quality table for the first stage only. Interpret the coefficient on your instrument from the first stage.**

In the first stage equation in IV design, researchers examine the effect of the instrumental variable (i.e., nudge; z; here, Republican judge) on compliance (i.e., get a longer sentence under a Republican judge). According to Table 2, it turned out that, on average, criminals spend 3.2 months more in prison when they are assigned to Republican judges compared to Democrat judges while controlling for the severity of crime. Therefore, the effect of nudge on compliance is confirmed.

Table 2. First stage regression

	(1)
	Months in jail
Republican judge	3.222^{***}
	(8.77)

Severity of crime	18.15***
	(80.21)
_cons	-19.47***
	(-37.46)
N	5000

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- **Calculate the "reduced form." Calculate the ratio of the reduced form.**

I got a coefficient of 0.14 for Republican judge (B_{RF}) after running reduced from regression (IV: Republican judge, DV: Recidivates, Control: Severity of crime). For B_{IV}, I divided B_{RF} by B from the first regression and got 0.14/3.22=0.04. I would claim the treatment effect of months in jail on recidivates is 0.04 based on the 2SLS results.

**Results from reduced form regression

```
. reg recidivates republicanjudge severityofcrime

      Source |       SS           df       MS      Number of obs   =     5,000
-------------+----------------------------------   F(2, 4997)      =    366.23
       Model |  141.631142         2  70.8155709   Prob > F        =    0.0000
    Residual |  966.239058     4,997   .19336383   R-squared       =    0.1278
-------------+----------------------------------   Adj R-squared   =    0.1275
       Total |   1107.8702     4,999  .221618364   Root MSE        =    .43973

------------------------------------------------------------------------------
 recidivates | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
republicanjudge |  .1426641   .0124434    11.47   0.000     .1182697    .1670586
severityofcrime |  .1885599   .0076618    24.61   0.000     .1735395    .2035803
       _cons | -.1139591   .0175983    -6.48   0.000    -.1484595   -.0794587
------------------------------------------------------------------------------
```

- **Now complete the IV regression and make a publication quality table of the second stage. Compare your answer to question #8 (above) to the IV coefficient in #9. State the F-stat in your writeup. It does not need to go into your table (although, in an actual publication it would). Is it above the conventional threshold?**

I run IV regression using ivreg2 command. Table 3 shows the results from the IV regression. The coefficient of Months in jail (x) indicates the effect treatment effect of months in jail on recidivates based on IV regression. It turned out that IV coefficient (i.e., coefficient of months in jail: 0.04) and the ratio of the reduced form (B_{RF}/B_{FS}) are the same. This is because IV coefficient is calculated by (B_{RF}/B_{FS}), the same ratio. The F-stat for the IV regression is 164.34 with the p-value<0.00. This is above the conventional threshold (around 10 or higher) and may signal the strength of the instrument.

Table 3. IV regression

	(1)
	Recidivates
Months in jail	0.0443***
	(7.68)
Severity of crime	-0.615***
	(-5.85)
_cons	0.748***
	(7.10)
N	5000

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- **Complete these sentences.**

In the research design above (using randomized judges), the ***always-takers*** *are the defendants who are always getting a longer sentence regardless of the assigned judge's political stance.*

The ***never-takers*** *are the defendants who are always getting a shorter sentence regardless of the assigned judge's political stance.*

The ***compliers*** *are the defendants who are getting a longer sentence only if the assigned judge is a Republican.*

The ***defiers*** *are the defendants who are getting a longer sentence only if the assigned judge is a Democrat.*

- **Comment on the monotonicity assumption and the possibility of "defiers" in this setting.**

I would say the monotonicity assumption is violated in this research. An instrumental variable approach assumes that there are no defiers. However, since the judge's political stance is not the only predictor of judge's decision, there can be Republican judges who less harshly sentence criminals than Democrat judges and vice versa. Thus, the possibility of defiers is high in this setting, considering some cases in which Democrat judges may harshly sentence criminals who engaged in hate crimes against social minorities.

- **In your dataset, what types of defendants are compliers?**

The compliers are the defendants who end up getting a longer sentence under a Republican judge and a shorter sentence under a Democrat judge.

- **Does the cycle of crime hypothesis appear to be true for the compliers?**

The cycle of crime hypothesis appears to be true for the compliers in this dataset. By adopting the judge's political stance as an instrumental variable and assuming the judge is randomly assigned, the effect of months in jail positively and statistically significantly affected recidivism.