

# Causal inference 25, Homework-1

Popov Platon

March 27, 2025

## 1 Homework # 1

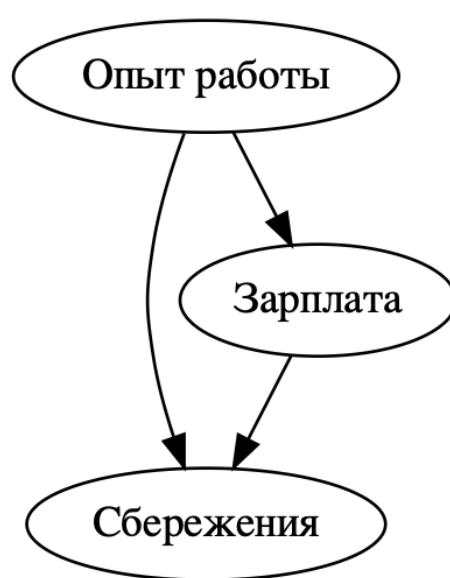
1. Explain the meme above.

*Мем заключается в игре слов, так как для приготовления коктейля пина колада требуется кокосы и ананасы, что соответствует типу графовой причинно-следственной структуре - collider, когда на одну переменную влияют сразу две других.*

2. Give an example (from applied economics, it could be labor, health, crime, education, trade etc.) of a causal structure that this meme refers to. Explain what problem can result from a straightforward application of OLS method to your example.

*Возьмем следующую причинно-следственную структуру и оценим влияние опыта работы на зарплату. В ней опыт работы влияет на зарплату: чем дольше человек работает, тем больше шанс, что он станет профессионалом своего дела, которым платят больше остальных. Кроме того, опыт работы влияет на уровень сбережений: чем дольше человек откладывает с доходов от профессиональной деятельности, тем больше у него сбережений. Однако, зарплата также влияет на сбережения: чем больше зарплата, тем больше человек может сберегать. Таким образом, дополнительно учитываемые, сбережения являются коллаيدر.*

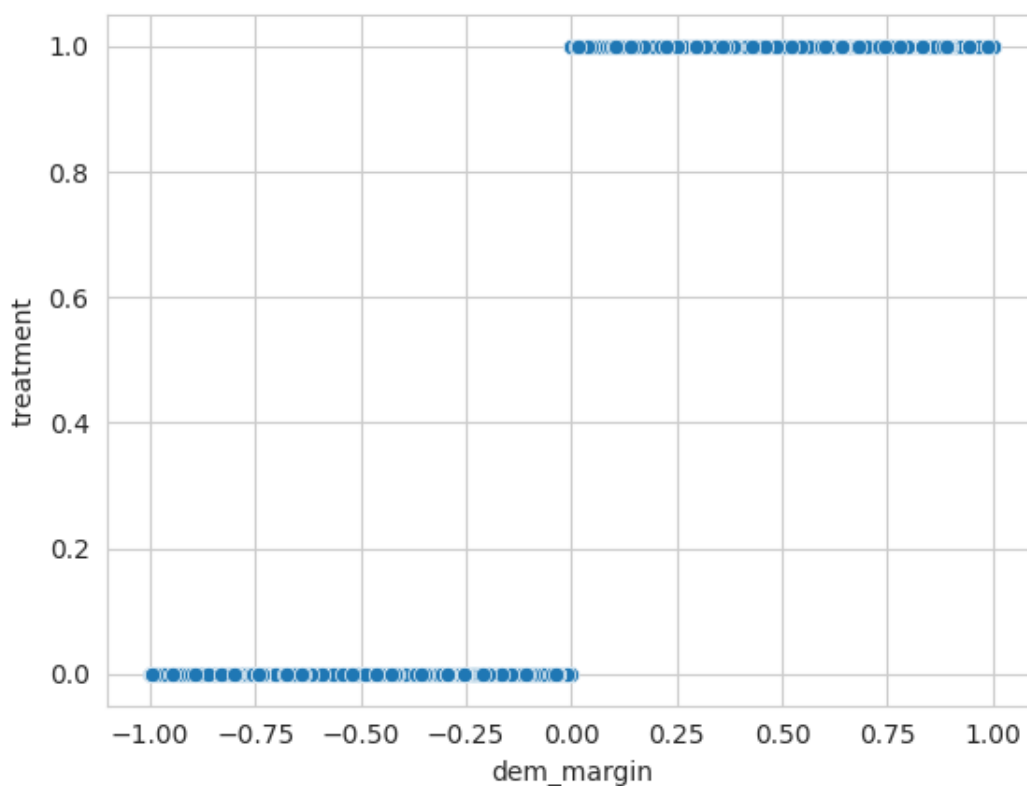
*Потенциально может возникнуть проблема с применением OLS, если мы проконтролируем коллаيدر, тогда возникнет смещение оценки влияния опыта работы на зарплату. Например, мы возьмем только тех, у кого накоплено большое количество сбережений, тогда оценка влияния опыта работы будет смещена, так как люди с большими сбережениями независимо от опыта работы, скорее будут иметь большую зарплату.*



4. Explain, both mathematically and substantively, what the key assumptions needed to make causal inferences in this setting. Is this a sharp RD design or a fuzzy RD design?

*В статье оценивается локальный эффект преимущества партии власти на следующих выборах. Преимущество оценивается, как разница между долями голосов. Математически нужно оценить значимость коэффициента при дамми-переменной, отвечающей за *threshold*, которая 1 если разница  $> 0$ , 0 если меньше. Если коэффициент значим, то победа на предыдущих выборах, действительно, влияет на результаты будущих выборов. В данном случае используется *sharp RD design*, поскольку оценивается именно влияние на победу, а не на то, займет кандидат должность или нет, для чего мог бы быть применен *fuzzy RD design*, если кандидат после победы отказался бы от поста.*

5. Generate the forcing variable and the treatment variables from the given data. Note that in its current form has rows for each candidate in a race. Plot the treatment as a function of the forcing variable to check the RD design.



*Похоже на то, что есть эффект от победы на прошлых выборах, поскольку разрыв оказывается именно около нуля.*

6. For the Democrats, run a model of the Democratic vote share in the next election on the forcing variable (Democratic margin of victory), allowing for different slopes for Democratic incumbents and nonincumbents. Report the coefficients and standard errors from this analysis and give a substantive interpretation of the coefficient on the treatment variable.

*Результаты регрессии из пакета *rdd* в питоне.*  
*Для кандидатов, победивших на прошлых выборах, доля голосов на следующих выборах увеличивается на 0.077 процентных пункта.*

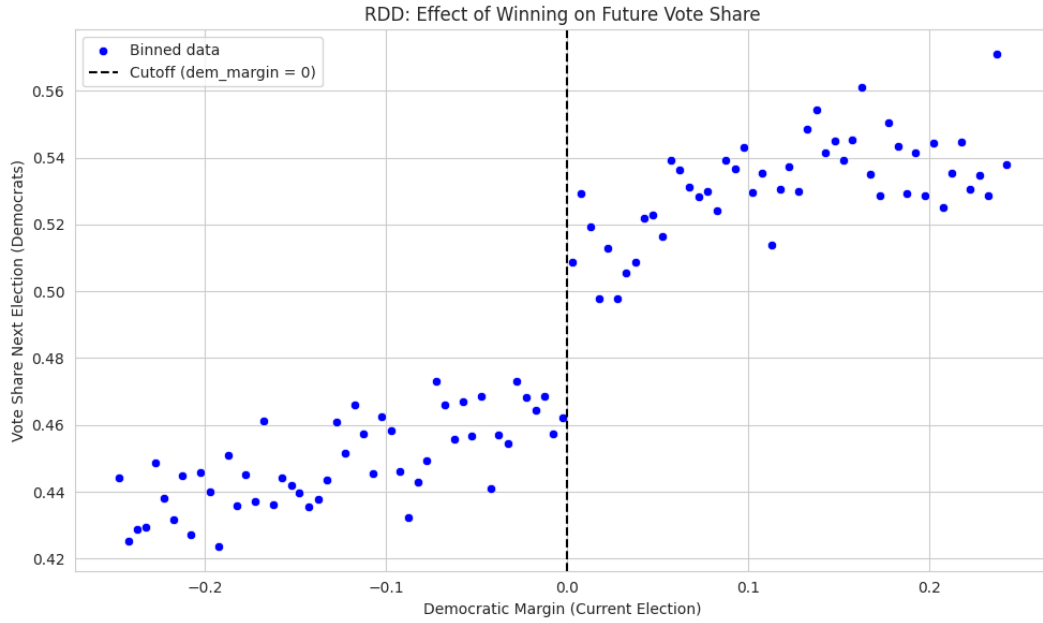
Table 1: Regression Discontinuity Design Results with Interaction Term

	Dependent Variable: Democrat Vote Share (t+1)		
	Coefficient	Std. Error	t-value
Intercept	0.470***	(0.002)	235.00
Democratic Margin (t)	0.336***	(0.006)	56.00
Threshold	0.077***	(0.004)	19.25
Margin $\times$ Threshold	-0.001	(0.008)	-0.13

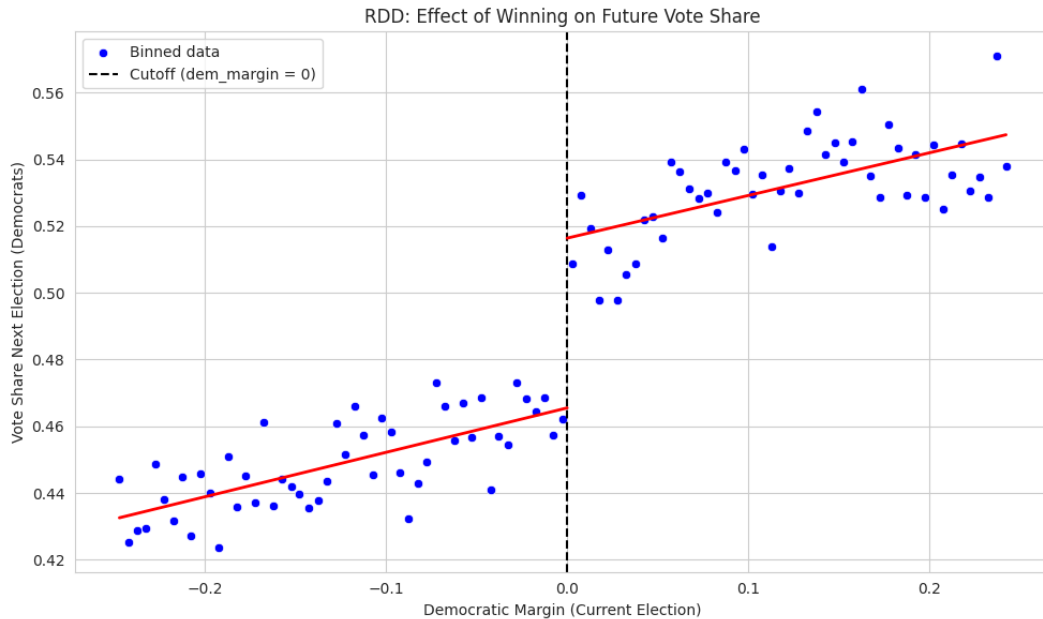
Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Robust standard errors in parentheses.

Threshold variable equals 1 when Democratic Margin  $> 0$ .

The interaction term tests for differential slopes on either side of the cutoff.

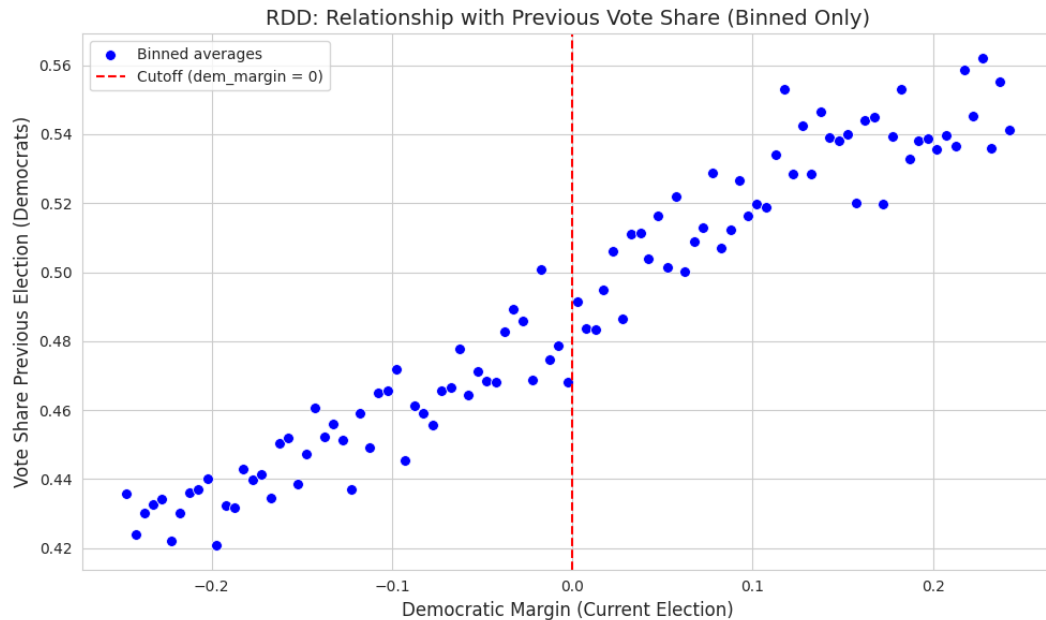


7. Plot the fitted values from this analysis over a scatterplot of the relationship between the forcing and outcome variables.



8. Perform a placebo test of the effect by replacing the next election's Democratic vote share with the previous election's Democratic vote share. Do you see any potential violations of the identifying assumptions?

*Проблема идентификации заключается в том, что сложно использовать нахождение у власти сегодня для того, чтобы получить больше голосов несколько лет назад. Как видно на картинке, значимой разницы между двумя периодами нет.*



9. Perform a placebo test of the effect by replacing the next election's Democratic vote share with the number of previous victories the candidate has won (officeexp). Do you see any potential violations of the identifying assumptions?

*Та же самая проблема: выиграть в прошлом невозможно засчет будущих голосов. На картинке сложно заметить прыжок в регрессии.*

