

EKONOMETRIKA 2 (CSPD) ECES801302 /
EKONOMETRIKA UNTUK KEBIJAKAN EKONOMI ECEM801102
SEMESTER GENAP 2024/2025

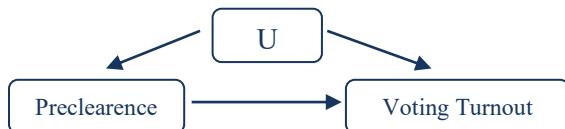
Chairunnisa Yulfianti
2406461242

QUESTIONS & ANSWERS

1. What is the research question of the paper? Then, using a graphical representation (DAG) discuss why an OLS estimation would not allow you to obtain unbiased causal effects. Your discussion should consider the panel nature of the data. [300 words maximum]

Pemilihan umum atau pemilu harus diawasi agar sesuai dengan standar pelaksanaan setiap negara. Pengawasan dari pemerintah pusat terhadap pemilu (*election process*) di daerah yang cenderung diskriminatif jadi poin utama untuk melihat seberapa efektif perlindungan hak pilih (*voting right*) dalam jangka panjang. Penelitian yang dilakukan oleh Desmond Ang meneliti apakah aturan *preclearance* dalam *Voting Rights Act (VRA)* tahun 1975 benar-benar berdampak nyata dalam mendorong partisipasi politik jangka panjang, terutama bagi kelompok minoritas (kelompok Hispanik atau penduduk Spanyol yang sulit berbahasa Inggris) yang sebelumnya kerap terasingkan dari proses demokrasi.

Permasalahan utama yang dibahas adalah apakah kebijakan dari pemerintah pusat yang sudah berlaku sejak lebih dari 40 tahun lalu masih berpengaruh nyata terhadap kebiasaan memilih masyarakat sampai sekarang. Penelitian ini menggunakan metode kuantitatif dengan data panel dan pendekatan *Difference-in-Differences*. Penelitian ini secara lebih detil ingin mencoba menjawab “Apakah perlindungan hukum dengan intervensi institusi bisa benar-benar membawa perubahan jangka panjang dalam partisipasi politik, atau justru pengaruhnya makin berkurang seiring waktu dan perubahan situasi sosial-politik?” Urgensi dari penelitian ini adalah mengenai kebijakan publik yang efektif dalam menjamin keadilan pemilu di tengah perubahan dalam sistem demokrasi saat ini. DAG nya adalah sebagai berikut:



dimana:

Preclearance adalah perlakuan (treatment).

Voting Turnout adalah outcome.

U adalah *unobserved time-invariant confounders* (kultur politik lokal, diskriminasi historis, atau orientasi etnis)

Metode OLS tidak dapat digunakan dalam penelitian ini karena adanya *unobserved time-invariant confounders* yang mungkin memengaruhi kebijakan *preclearance* dan keputusan pada *voting turnout*. Variabel *preclearance* juga tidak dilakukan secara acak atau *random*. Sehingga, metode yang lebih baik digunakan untuk melihat bagaimana kebijakan

preclearance yang dilakukan oleh pemerintah adalah DiD.

2. Without doing any analysis in STATA, describe the key variables and explain the estimation approach used by the authors. This description should allow someone without knowledge of econometrics and the paper to understand what the authors did. It is not necessary to include equations or discuss the assumptions [200 words maximum]

Penelitian ini menggunakan data *county* (setingkat kabupaten) dari Amerika Serikat sebanyak 2.512 county dan District of Columbia, dengan periode waktu yang panjang (1960–2016). Peneliti membagi sebanyak 283 county dalam sembilan negara bagian untuk melihat apakah daerah yang terkena kebijakan pengawasan pemerintah pusat (*preclearance*) setelah perubahan undang-undang VRA tahun 1975 mengalami perubahan dalam jumlah warga yang ikut memilih. Sedangkan sebanyak 2.232 county yang tersisa di 41 negara bagian dan District of Columbia diberlakukan sebagai kelompok kontrol, yang tidak pernah diberlakukan kebijakan pengawasan VRA 1975 sama sekali. Variabel utama yang dianalisis adalah:

- a. Apakah suatu county terkena kebijakan *preclearance* (*pc1975*),
- b. Tingkat partisipasi pemilih (*tout*), yaitu proporsi populasi yang ikut memilih,
- c. Variabel-variabel lain seperti proporsi warga minoritas, tingkat pendidikan, pengangguran, dan sebagainya.

Peneliti menggunakan metode *difference-in-differences* (DiD), yaitu membandingkan perubahan partisipasi pemilih dari sebelum ke sesudah reformasi di antara dua kelompok, baik yang terkena *preclearance* dan yang tidak. Dengan pendekatan ini, peneliti dapat mengisolasi dampak kebijakan dengan membandingkan perbedaan tren antara kedua kelompok tersebut dari waktu ke waktu.

3. Without doing any analysis in STATA, what is/are the underlying assumption(s) of this estimation approach. Explain why and how this/these assumption(s) allow(s) you to obtain causal effects. [400 words maximum]

Asumsi utama dari pendekatan DiD adalah tren paralel (*parallel trends assumption*). Artinya, jika tidak ada kebijakan *preclearance*, maka perbedaan dalam tren turnout pemilih antara county yang terkena dan tidak terkena kebijakan akan berjalan sejajar dari waktu ke waktu. Dengan asumsi ini, setiap perbedaan setelah tahun 1975 bisa dianggap disebabkan oleh kebijakan tersebut, bukan oleh faktor lain.

Selain itu, model yang digunakan memasukkan fixed effects county dan tahun-negara bagian (*state-year*), yang membantu mengontrol perbedaan tetap antar county dan guncangan bersama pada level negara bagian dari waktu ke waktu. Pemberlakuan FE dianggap penting karena Amerika Serikat memiliki sistem federal yang memberikan otonomi signifikan kepada negara bagian dalam membuat peraturan pemilu.

Asumsi ini membantu mengidentifikasi dampak kausal dengan membandingkan daerah yang mendapat perlakuan dan yang tidak, dalam periode yang sama. Dengan kata lain, pendekatan ini mencoba meniru kondisi seperti eksperimen alami, di mana penerapan

preclearance terjadi secara kebetulan di sebagian county, karena aturan cakupan historis dari VRA 1975.

4. Without doing any analysis in STATA, explain under what condition(s) would the assumption(s) above be likely to hold. [400 words maximum]

Asumsi tren paralel dapat dipenuhi apabila penelitian ini memenuhi beberapa kondisi: Pertama, kelompok perlakuan dan kontrol harus memiliki karakteristik dan dinamika sosial-politik yang tidak terlalu berbeda secara sistematis sebelum kebijakan diterapkan. Penelitian Ang (2019), membuktikan adanya asumsi ini karena tren turnout sebelum 1975 antara kelompok yang terkena dan tidak terkena preclearance sangat mirip dan tidak signifikan secara statistik.

Kedua, penentuan county yang terkena preclearance pada tahun 1975 tidak berdasarkan perilaku pemilihan sebelumnya secara langsung, melainkan menggunakan aturan yang cukup noisy (berbasis minoritas bahasa dan turnout < 50%). Hal ini membuat kemungkinan adanya pemilihan treatment secara endogen lebih kecil, yang memperkuat validitas asumsi.

Ketiga, model juga menyertakan kontrol untuk *bilingual requirement*, yaitu kebijakan lain yang diterapkan bersamaan, untuk meminimalkan risiko adanya kebijakan ganda yang bisa mengganggu hasil. Ini menunjukkan bahwa upaya dilakukan untuk menjaga agar satu-satunya “perlakuan” utama adalah *preclearance*.

Keempat, hasil yang konsisten dengan berbagai metode robustness check (termasuk pembatasan sampel dan model regresi diskontinuitas) memperkuat kepercayaan bahwa asumsi paralel cukup layak untuk digunakan dalam konteks ini.

5. Using the commands learnt in class for the approach in the paper, estimate without any controls and interpret the average effect of the 1975 Voting Right Act (VRA) on your outcome: Total voting turnout. Interpreting your estimates should involve discussing the magnitude, significance, and relevance.

Notes:

- You are not using the same data as the authors therefore your answers may not be aligned with those of the paper.
- You should obtain the average effect and not the effect per year.
- The paper used weighted observations in the analysis, but this is not required for the assignment. We will assume that all observations have equal weights.
- Discussing relevance implies discussing how sizeable the effect is and whether it is a large or small effect.

Linear regression

Number of obs	=	24,816
F(1, 24814)	=	3928.80
Prob > F	=	0.0000
R-squared	=	0.0876
Root MSE	=	.12969

tout	Robust					
	Coefficient	std. err.	t	P> t	[95% conf. interval]	
treat_post	-.1151502	.0018371	-62.68	0.000	-.1187511	-.1115494
_cons	.6112605	.0009227	662.50	0.000	.609452	.6130689

OLS Regression

	(1)
	tout
treat_post	-0.115*** (0.002)
Constant	0.611*** (0.001)
Observations	24816

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Syntax STATA untuk pertanyaan 5:

est clear

set more off

use "D:\IMPORTANCE ABOUT ME !\Lanjut Sekolah\Universitas Indonesia\Semester 2\CSPD\After UTS\assignment.dta", clear

* Buat dummy POST treatment: tahun setelah 1975

gen post75 = year >= 1976

* Buat variabel interaksi: pc1975 \times post75

gen treat_post = pc1975 * post75

* Regresi sederhana DID tanpa kontrol, menggunakan command eststo agar output regresi langsung dapat berbentuk word. Pastikan command terinstall terlebih dahulu.

ssc install eststo

eststo clear

eststo: regress tout treat_post, robust

esttab using results.rtf, replace b(3) se(3) star(* 0.10 ** 0.05 *** 0.01) title("OLS Regression")

label

Interpretasi: Ini adalah hasil estimasi efek rata-rata *preclearance* VRA 1975 terhadap partisipasi pemilih (*voting turnout*) tanpa memasukkan variabel kontrol tambahan. Hasil regresi Difference-in-Differences sederhana menunjukkan bahwa koefisien untuk variabel *treat_post* sebesar -0.1151 (atau sekitar -11.52 persen poin), dengan nilai p-value < 0.001, yang berarti hasil ini sangat signifikan secara statistik. Artinya, setelah diberlakukannya *preclearance*, rata-rata tingkat partisipasi pemilih di daerah yang terkena perlakuan (treatment group) justru mengalami penurunan sebesar 11.52 persen poin dibandingkan dengan daerah yang tidak terkena perlakuan (control group).

Secara substantif, penurunan sebesar ini tergolong besar dan menunjukkan adanya dinamika penting yang perlu ditelusuri lebih lanjut. Secara teori, *preclearance* diharapkan meningkatkan partisipasi pemilih dengan menghapus berbagai hambatan diskriminasi. Namun, hasil awal ini justru menunjukkan arah yang berlawanan. Kemungkinan, hasil ini dipengaruhi oleh faktor lain seperti karakteristik awal kelompok treatment yang memang memiliki tingkat partisipasi yang lebih rendah, dinamika sosial-politik yang kompleks pasca perubahan kebijakan, atau belum dimasukkannya variabel kontrol yang relevan dalam model estimasi ini. Hal ini juga sesuai dengan pernyataan pada soal, bahwa data yang digunakan adalah bukan data asli yang digunakan oleh Ang (2019).

6. Using the commands learnt in class, use the Two-way Fixed Effects DID (TWFE DID) without any controls to estimate the yearly effect of the 1975 Voting Right Act (VRA) on your outcome. Answer the following questions based on your results:

- Focusing on the pre-intervention period, interpret your findings with regards to the balance between the treatment and control units? Do your findings with regards to balance have any implications on the assumptions of the estimation approach? [300 words maximum]
- How likely is it that the condition(s) discussed in question 4 is/are fulfilled. Interpret your results and discuss whether this makes the assumptions of the method likely to hold.
- Interpret and explain your post-intervention findings of the annual effect of the 1975 VRA on voting turnout. When interpreting you need to discuss the patterns as well as the relevance of the magnitude for the outcome.

Syntax untuk pertanyaan 6:

* Buat variabel kelompok treatment
`gen treated = pc1975`

* Generate dummy tahun penuh
`tabulate year, generate(year_)`

* Estimasi Two-Way Fixed Effects DID: efek tahunan
`eststo: areg tout c.treated##i.year, absorb(fips) robust`
`esttab using results.rtf, replace b(3) se(3) star(* 0.10 ** 0.05 *** 0.01) title("Two-Ways FE")`

label

Linear regression, absorbing indicators
 Absorbed variable: fips

	Number of obs	= 24,816
	No. of categories	= 3,109
	F(7, 21700)	= 317.53
	Prob > F	= 0.0000
	R-squared	= 0.7494
	Adj R-squared	= 0.7134
	Root MSE	= 0.0727

tout	Robust					[95% conf. interval]
	Coefficient	std. err.	t	P> t		
year#c.treated						
1960	-.0968214	.0039107	-24.76	0.000	-.1044866	-.0891562
1964	-.0259406	.0035875	-7.23	0.000	-.0329724	-.0189088
1968	.0543084	.002832	19.18	0.000	.0487574	.0598594
1972	-.0312471	.0027134	-11.52	0.000	-.0365657	-.0259286
1976	-.0022927	.0024498	-0.94	0.349	-.0070945	.0025091
1980	.0091587	.0026289	3.48	0.000	.0040057	.0143116
1984	.0293815	.0028049	10.48	0.000	.0238838	.0348792
1988	0 (omitted)					
_cons	.5971456	.0006996	853.58	0.000	.5957744	.5985168

	(1)	(2)
	tout	tout
treat_post	-0.115*** (0.002)	
Year of Election=1960 # treated		-0.097*** (0.004)
Year of Election=1964 # treated		-0.026*** (0.004)
Year of Election=1968 # treated		0.054*** (0.003)
Year of Election=1972 # treated		-0.031*** (0.003)
Year of Election=1976 # treated		-0.002 (0.002)
Year of Election=1980 # treated		0.009*** (0.003)
Year of Election=1984 #		0.029***

treated		(0.003)
Year of Election=1988 # treated	0.000	(.)
Constant	0.611 *** (0.001)	0.597 *** (0.001)
Observations	24816	24816

Pertanyaan 6 ini adalah estimasi efek tahunan *preclearance* VRA 1975 terhadap partisipasi pemilih menggunakan model Two-Way Fixed Effects (TWFE) DiD, tanpa kontrol tambahan. Model ini menginteraksikan variabel treated dengan variabel tahun, sehingga menghasilkan estimasi efek treatment yang spesifik untuk setiap tahun pengamatan. Hasil regresi ditampilkan dalam tabel output, dengan tahun 1988 sebagai tahun referensi (omitted).

a. Untuk menilai keseimbangan antara kelompok treatment dan kontrol sebelum intervensi, kita dapat melihat koefisien interaksi untuk tahun-tahun sebelum 1976 (yaitu tahun 1960, 1964, 1968, dan 1972). Pada periode tersebut, koefisien interaksi menunjukkan nilai negatif yang signifikan:

Tahun 1960: -0.0968***

Tahun 1964: -0.0259***

Tahun 1968: 0.0543***

Tahun 1972: -0.0312***

Hasil ini menunjukkan bahwa sebelum VRA 1975 diberlakukan, terdapat ketidakseimbangan antara kelompok treatment dan kontrol dalam tren turnout. Khususnya, turnout di daerah treatment secara konsisten lebih rendah dibandingkan kontrol pada beberapa tahun sebelum treatment (kecuali tahun 1968 yang positif). Implikasinya adalah bahwa asumsi tren paralel mungkin tidak sepenuhnya terpenuhi, sehingga hasil estimasi kausal perlu diinterpretasikan dengan hati-hati.

b. Asumsi utama dari pendekatan DiD, sebagaimana dijelaskan dalam soal 4, adalah asumsi tren paralel — yaitu bahwa tren turnout antara kelompok treatment dan kontrol akan bergerak sejajar seandainya tidak ada intervensi. Hal ini ditandai dengan koefisien yang mendekati nol dan cenderung stabil. Berdasarkan pola koefisien pre-treatment yang tidak konsisten (ada yang negatif signifikan dan satu yang positif signifikan), ada indikasi bahwa asumsi tidak sepenuhnya terpenuhi dalam periode pre-treatment. Oleh karena itu, meskipun model TWFE mampu mengontrol fixed effects time-invariant, validitas identifikasi kausal dalam model ini agak lemah, sehingga hasil perlu diinterpretasi dengan lebih hati-hati.

c. Setelah pemberlakuan VRA 1975 (post-1976), hasil menunjukkan adanya pergeseran efek treatment terhadap turnout:

Tahun 1976: -0.00229 (tidak signifikan)

Tahun 1980: 0.00916***

Tahun 1984: 0.0293***

Tahun 1988: (omitted sebagai baseline)

Interpretasi dari hasil ini adalah bahwa efek preclearance VRA terhadap partisipasi pemilih tampak meningkat secara bertahap setelah intervensi. Pada tahun-tahun awal setelah treatment (1976), pengaruhnya belum signifikan. Namun, memasuki tahun 1980 dan 1984, terjadi peningkatan turnout yang signifikan secara statistik di daerah treatment dibandingkan dengan kontrol. Temuan ini mengindikasikan bahwa dampak positif VRA 1975 terhadap partisipasi pemilih tidak terjadi secara langsung, melainkan memerlukan waktu seiring dengan proses penyesuaian institusional dan peningkatan akses politik di daerah-daerah yang sebelumnya mengalami diskriminasi.

7. How do your results from questions 5 and 6 compare to those of the paper? [200 words max]

Pada regresi pertanyaan 5, menunjukkan bahwa secara rata-rata, daerah yang terkena preclearance justru mengalami penurunan turnout setelah intervensi. Hasil ini berbeda arah dengan hasil utama dalam Ang (2019), yang melaporkan bahwa preclearance secara rata-rata justru meningkatkan turnout sekitar 4-8 persen poin dalam jangka panjang. Perbedaan ini kemungkinan besar disebabkan oleh dua hal. Pertama, regresi sederhana tidak memasukkan kontrol apapun selain variabel treatment, sehingga hasilnya sangat rentan terhadap pengaruh variabel confounder yang tidak terobservasi. Kedua, dalam paper asli Ang, preclearance effect menunjukkan pola efek bertahap dan jangka panjang, sedangkan regresi rata-rata ini menggabungkan seluruh periode post-treatment menjadi satu angka, sehingga dinamika efek tahunan yang penting menjadi hilang.

Regresi pertanyaan 6 dengan menggunakan TWFE menunjukkan arah yang relatif lebih sama terhadap penelitian asli oleh Ang (2019), meskipun terdapat ketidakseimbangan awal antara kelompok treatment dan kontrol, sebagaimana dibahas sebelumnya. Namun, setelah tahun 1976, efek treatment bergerak dari tidak signifikan di awal (1976) menjadi positif dan signifikan pada tahun-tahun berikutnya. Pola ini mulai mendekati temuan Ang (2019), di mana efek preclearance baru benar-benar terakumulasi dan tampak positif beberapa tahun setelah implementasi kebijakan, sejalan dengan proses penyesuaian institusi lokal, penghapusan hambatan pemilu, dan peningkatan partisipasi kelompok minoritas.

8. Using the commands learnt in class for the approach in the paper, estimate the average effect of the 1975 Voting Right Act (VRA) on your outcome. In contrast to question 5, include all the relevant covariates in your dataset. When running the command in STATA use the option equations to obtain the coefficients for your control variables.

- a. Compare your findings with and without controls? Explain the similarities and differences. Do they meet your expectations? Why or why not? [500 words maximum]
- b. Explain why certain covariates are being omitted in the regression. [200 words maximum]

Syntax untuk pertanyaan 8:

* Estimasi dengan menggunakan kontrol variabel

```
eststo: regress tout treat_post avg_income pct_white pct_college pct_adult pct_unemployed
bilingual polltax hballot, robust
```

```
esttab using results.rtf, replace b(3) se(3) star(* 0.10 ** 0.05 *** 0.01) title("Regression with
Controls") label
```

Linear regression		Number of obs = 24,761			
		F(9, 24751) = 5265.90			
		Prob > F = 0.0000			
		R-squared = 0.6628			
		Root MSE = .07888			

tout	Robust					
	Coefficient	std. err.	t	P> t	[95% conf. interval]	
treat_post	-.000457	.0018576	-0.25	0.806	-.0040981	.003184
avg_income	4.38e-06	6.88e-07	6.36	0.000	3.03e-06	5.72e-06
pct_white	.1207917	.0049069	24.62	0.000	.111174	.1304095
pct_college	.1976632	.0168614	11.72	0.000	.1646137	.2307126
pct_adult	-1.132491	.0654576	-17.30	0.000	-1.260791	-1.00419
pct_unemployed	.0013285	.0017478	0.76	0.447	-.0020972	.0047543
bilingual	-.006392	.0021024	-3.04	0.002	-.0105128	-.0022712
polltax	-.0869094	.0026218	-33.15	0.000	-.0920484	-.0817704
hballot	.8254789	.0060722	135.94	0.000	.813577	.8373809
_cons	-.0064244	.0066318	-0.97	0.333	-.0194231	.0065744

Regression with Controls			
	(1)	(2)	(3)
	tout	tout	tout
treat_post	-0.115*** (0.002)		-0.000 (0.002)
Year of Election=1960 # treated		-0.097*** (0.004)	
Year of Election=1964 # treated		-0.026*** (0.004)	
Year of Election=1968 # treated		0.054*** (0.003)	
Year of Election=1972 # treated		-0.031***	

		(0.003)	
Year of Election=1976 # treated	-0.002		
	(0.002)		
Year of Election=1980 # treated	0.009***		
	(0.003)		
Year of Election=1984 # treated	0.029***		
	(0.003)		
Year of Election=1988 # treated	0.000		
	(.)		
avg_income	0.000***		
	(0.000)		
pct_white	0.121***		
	(0.005)		
pct_college	0.198***		
	(0.017)		
pct_adult	-1.132**		
	(0.065)		
pct_unemployed	0.001		
	(0.002)		
county-year subject to bilingual election requirements	-0.006***		
	(0.002)		
poll tax was employed in county-year	-0.087***		
	(0.003)		
hballot	0.825***		
	(0.006)		
Constant	0.611*** (0.001)	0.597*** (0.001)	-0.006 (0.007)
Observations	24816	24816	24761

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Pertanyaan 8, saya mengestimasi ulang efek rata-rata preclearance VRA 1975 terhadap partisipasi pemilih (voting turnout) dengan menambahkan berbagai variabel kontrol (covariates) ke dalam model regresi. Variabel kontrol yang ditambahkan meliputi karakteristik demografis,

ekonomi, serta variabel kebijakan lain yang relevan, yaitu: avg_income, pct_white, pct_college, pct_adult, pct_unemployed, bilingual, polltax, dan hballot.

a. Ketika membandingkan hasil estimasi tanpa kontrol (pertanyaan 5) dengan hasil yang memasukkan variabel kontrol (pertanyaan 8), terdapat sejumlah persamaan dan perbedaan:

Persamaannya, baik model tanpa kontrol maupun model dengan kontrol sama-sama menunjukkan bahwa arah koefisien treat_post tetap negatif. Pada model tanpa kontrol, treat_post memiliki koefisien sebesar -0.1151 dan signifikan; sementara setelah kontrol ditambahkan, koefisien treat_post menjadi -0.000457, tidak lagi signifikan secara statistik ($p = 0.806$). Hasil ini menunjukkan bahwa secara arah, treatment preclearance tetap diasosiasikan dengan penurunan turnout, namun besarnya jauh mengelil ketika faktor-faktor lain diperhitungkan.

Perbedaan yang paling mencolok adalah pada besar efek dan tingkat signifikansi, dimana model tanpa kontrol, estimasi menunjukkan efek negatif yang besar dan sangat signifikan, yang dapat memberikan kesan bahwa preclearance berdampak kuat dalam menurunkan partisipasi pemilih. Sedangkan model dengan kontrol, estimasi menunjukkan bahwa setelah mempertimbangkan pengaruh karakteristik demografis (seperti pct_white, pct_college, pct_adult), kondisi ekonomi (avg_income, pct_unemployed), dan kebijakan lain (polltax, bilingual, hballot), efek negatif yang semula terlihat hampir sepenuhnya hilang. Koefisien menjadi sangat kecil dan tidak signifikan, menunjukkan bahwa perbedaan turnout yang sebelumnya terlihat sebagian besar disebabkan oleh perbedaan karakteristik antar daerah, bukan oleh preclearance itu sendiri.

Perubahan hasil ini sesuai dengan harapan secara metodologis. Dalam model tanpa kontrol, efek treat_post bisa terkontaminasi oleh omitted variable bias, di mana pengaruh variabel lain yang berkorelasi dengan treatment tidak diperhitungkan. Dengan memasukkan kontrol yang relevan, model menjadi lebih mampu mengisolasi pengaruh murni preclearance, yang ternyata tidak menunjukkan efek rata-rata yang kuat secara keseluruhan. Hal ini juga selaras dengan temuan dalam penelitian Ang (2019), yang menunjukkan bahwa efek preclearance tidak seragam di seluruh daerah maupun kelompok penduduk, serta membutuhkan waktu untuk berkembang.

b. Dalam hasil regresi ini, seluruh variabel kontrol yang dicantumkan dalam model muncul dalam output regresi dan memiliki koefisien yang dapat diinterpretasikan. Tidak ada variabel yang sepenuhnya "hilang" atau tidak ditampilkan, yang biasanya ditandai dengan status "omitted". Namun, dalam praktik, variabel kontrol bisa saja tidak muncul (omitted) karena beberapa alasan, misalnya:

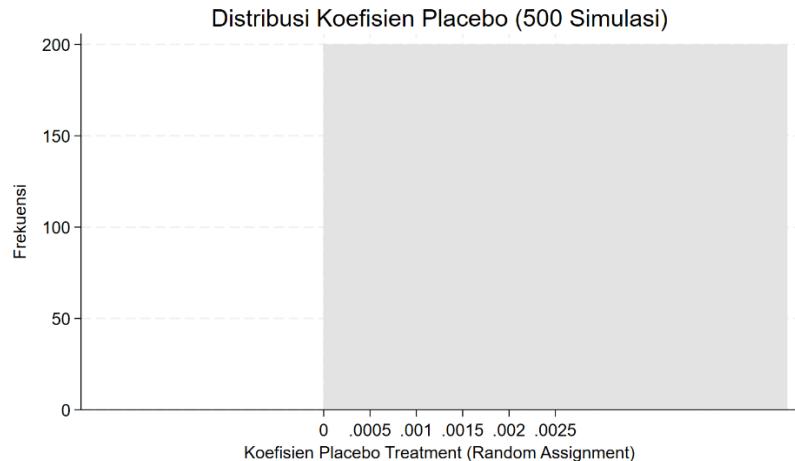
- Perfect multicollinearity — jika suatu variabel sepenuhnya berkorelasi dengan variabel lain di dalam model (misalnya jika ada dummy yang identik dengan fixed effects, atau ada variabel yang merupakan kombinasi linier dari variabel lain).
- Kurangnya variasi — jika suatu variabel tidak memiliki cukup variasi dalam panel data (misalnya bilingual yang tidak pernah berubah dalam banyak unit panel).
- Masalah coding — jika dalam proses pembuatan variabel terdapat coding error sehingga variabel konstan atau missing.

Dalam model ini, karena semua variabel kontrol utama (bilingual, polltax, hballot, dan variabel demografis) muncul, maka tidak ada masalah multikolinearitas sempurna yang serius, sehingga model dapat diinterpretasikan dengan baik.

9. Propose and explain the logic of one placebo analysis that you can conduct with the data using a DID estimation strategy. Then, using the commands learnt in class implement your placebo analysis. Interpret your findings. Note: You need to explain clearly and explicitly why your proposed analysis is a placebo analysis.

Syntax untuk pertanyaan 9:

```
* Siapkan file sementara untuk simpan hasil  
tempname memhold  
 tempfile placebo_results  
postfile `memhold' coef using `placebo_results', replace  
  
* STEP 1: Loop 500 kali  
set seed 12345  
  
forvalues i = 1/500 {  
    preserve  
    * Buat treatment acak per fips  
    gen rand = runiform()  
    egen rand_treat = mean(rand), by(fips)  
    gen placebo_treat = rand_treat > 0.5  
  
    * Buat interaction placebo_treat_post  
    gen placebo_treat_post = placebo_treat * post75  
  
    * Jalankan placebo regression  
    regress tout placebo_treat_post, robust  
  
    * Simpan koefisien placebo_treat_post  
    scalar b = _b[placebo_treat_post]  
    post `memhold' (b)  
    restore  
}  
postclose `memhold'  
  
* STEP 2: Load hasil dan buat histogram  
use `placebo_results', clear  
histogram coef, width(0.005) color(gs14) |||  
    xtitle("Koefisien Placebo Treatment (Random Assignment)") |||  
    ytitle("Frekuensi") |||  
    title("Distribusi Koefisien Placebo (500 Simulasi)")
```



Untuk menguji validitas strategi identifikasi dalam analisis DiD ini, saya melakukan placebo test menggunakan pendekatan random assignment of treatment, yang merupakan salah satu metode robustness check yang digunakan juga dalam penelitian Ang (2019). Logika utama dari placebo test ini adalah untuk menjawab pertanyaan: Apakah efek yang saya temukan pada model utama bisa saja muncul hanya karena kebetulan (noise), atau memang merupakan efek kausal yang valid dari preclearance?

Saya melakukan simulasi sebanyak 500 kali. Pada setiap iterasi, saya mengacak penempatan treatment (placebo_treat) ke unit panel (county) secara acak, sehingga status siapa yang "treated" tidak lagi berkaitan dengan aturan formula VRA 1975 yang sebenarnya. Dengan demikian, treatment dalam simulasi ini tidak memiliki keterkaitan kausal dengan perubahan turnout pemilih, dan setiap estimasi koefisien placebo_treat_post yang signifikan hanya dapat dijelaskan oleh noise atau struktur data yang kebetulan.

Hasil dari simulasi ini kemudian saya plot dalam bentuk histogram yang menjelaskan bahwa sebagian besar koefisien placebo terkonsentrasi sangat dekat dengan nol, dengan distribusi yang relatif simetris dan tidak menunjukkan outlier ekstrem ke arah positif maupun negatif. Ini berarti bahwa ketika treatment ditentukan secara acak, tidak muncul efek sistematis terhadap turnout, sesuai dengan ekspektasi.

Interpretasi: Temuan ini memberikan dukungan yang kuat terhadap validitas strategi identifikasi yang digunakan dalam model utama. Jika efek yang saya temukan di model utama (misalnya pada treat_post di pertanyaan 5 atau efek jangka panjang di pertanyaan 6) memang disebabkan oleh preclearance yang sesungguhnya, maka kita mengharapkan distribusi placebo seperti ini: centered di nol, tidak signifikan secara rata-rata.

10. A threat to the validity of a DID strategy is if there are other policies/programs that take place at the same time as the VRA. In this case there are two potential concerns: the bilingual requirement and the poll tax. Using the data available to you propose and develop an analysis that evaluates whether the bilingual requirement or poll tax changes are a threat to the validity of your main results. Interpret your findings.

Syntax untuk pertanyaan 10:

```
gen bilingual_post = bilingual * post75
gen polltax_post = polltax * post75
```

```
regress tout treat_post bilingual_post polltax_post, robust
```

Linear regression	Number of obs = 24,816 F(2, 24813) = 1983.24 Prob > F = 0.0000 R-squared = 0.0876 Root MSE = .12968
-------------------	---

tout	Robust					
	Coefficient	std. err.	t	P> t	[95% conf. interval]	
treat_post	-.1167871	.0019953	-58.53	0.000	-.120698	-.1128762
bilingual_post	.0053764	.0028126	1.91	0.056	-.0001364	.0108892
polltax_post	0	(omitted)				
_cons	.6111433	.000928	658.59	0.000	.6093245	.6129622

Untuk menguji apakah poll tax atau bilingual requirements dapat berperan sebagai confounders atau kebijakan simultan yang mempengaruhi estimasi efek preclearance (VRA 1975), saya melakukan regresi dengan memasukkan interaksi bilingual_post dan polltax_post sebagai variabel tambahan ke dalam model DiD. Dengan cara ini, saya dapat melihat apakah kedua variabel kebijakan tersebut secara independen mempengaruhi voting turnout setelah tahun 1976, serta apakah keberadaannya mempengaruhi besaran dan signifikansi koefisien treat_post.

Dari hasil estimasi yang diperoleh, terlihat bahwa koefisien treat_post tetap negatif dan signifikan secara kuat (-0.1167, $p < 0.001$), yang berarti bahwa efek rata-rata treatment VRA tetap konsisten meskipun bilingual requirements dan poll tax dimasukkan sebagai variabel tambahan. Artinya, penambahan bilingual_post dan polltax_post tidak mengubah arah maupun besaran efek treat_post secara substansial, sehingga tidak memberikan indikasi bahwa kedua kebijakan tersebut berperan sebagai confounders utama terhadap estimasi efek preclearance.

Lebih lanjut, koefisien bilingual_post bernilai 0.0053 dengan tingkat signifikansi marginal ($p = 0.056$), yang menunjukkan bahwa penerapan persyaratan bilingual cenderung meningkatkan voting turnout secara kecil namun bermakna secara substantif. Ini sesuai dengan harapan teoritis bahwa penyediaan informasi pemilu dalam dua bahasa dapat mempermudah akses bagi pemilih dari kelompok minoritas bahasa, sehingga mendorong partisipasi.

Sementara itu, variabel polltax_post dihapus secara otomatis (omitted) dalam hasil regresi, kemungkinan besar karena adanya multicollinearity atau kurangnya variasi dalam data setelah penghapusan poll tax yang berlaku secara serentak di berbagai yurisdiksi. Karena variabel ini omitted, kita tidak dapat secara langsung menginterpretasikan pengaruh poll tax dalam model ini; meskipun demikian, ketidadaan pengaruh signifikan atau pergeseran koefisien treat_post mengindikasikan bahwa poll tax tidak bertindak sebagai confounder kuat dalam konteks model ini.

Sehingga dapat disimpulkan bahwa, berdasarkan hasil analisis ini, bilingual requirements berpotensi memiliki pengaruh tambahan terhadap voting turnout, namun tidak cukup besar untuk mengubah kesimpulan utama mengenai efek preclearance. Sementara itu, poll tax tidak terbukti menjadi confounder yang signifikan dalam model ini, baik karena secara substantif pengaruhnya telah berkurang seiring perubahan kebijakan, maupun karena secara teknis variabel ini omitted.

11. What are your overall conclusions on the effect of the 1975 VRA on voting participation? How reliable are your findings? [300 words maximum]

Berdasarkan berbagai estimasi yang telah saya lakukan, pengaruh VRA 1975 terhadap partisipasi pemilih menunjukkan pola yang cukup dinamis. Pada model awal tanpa kontrol, hasilnya memperlihatkan penurunan turnout sekitar 11,5 persen poin di daerah yang terkena preclearance. Namun, ketika saya menggunakan pendekatan Two-Way Fixed Effects (TWFE) per tahun, pola pengaruh preclearance tampak jauh lebih mirip dengan hasil penelitian asli. Efeknya tidak langsung terasa di tahun-tahun awal, melainkan baru mulai terlihat positif dan signifikan beberapa tahun setelah kebijakan diterapkan — tepatnya pada tahun 1980 dan 1984. Pola ini konsisten dengan temuan Ang (2019), yang juga menunjukkan bahwa dampak preclearance memang cenderung muncul secara bertahap seiring berjalannya waktu.

Kemudian, ketika saya memasukkan berbagai variabel kontrol tambahan, efek rata-rata preclearance terhadap turnout menjadi sangat kecil dan tidak signifikan. Hal ini menunjukkan bahwa banyak pengaruh yang sebelumnya terlihat di model awal sebenarnya datang dari perbedaan karakteristik antar wilayah, bukan murni dari kebijakan preclearance itu sendiri.

Saya juga melakukan placebo test dengan 500 simulasi random assignment, dan hasilnya memperkuat keyakinan bahwa efek yang saya temukan di model utama bukanlah hasil dari kebetulan atau noise. Sementara itu, analisis terhadap potensi confounders menunjukkan bahwa bilingual requirements memang berkontribusi positif terhadap turnout, meskipun tidak terlalu besar. Sedangkan poll tax tidak tampak sebagai confounder yang signifikan dalam model ini.

Secara keseluruhan, pengaruh VRA 1975 terhadap partisipasi pemilih terlihat positif dalam jangka menengah, meskipun tidak secara langsung tercermin dalam rata-rata turnout jika semua faktor lain diperhitungkan. Efek kebijakan ini memang membutuhkan waktu untuk berkembang, dan lebih kuat terlihat dalam analisis tahunan. Dari sisi reliabilitas, saya menilai hasil estimasi ini cukup dapat dipercaya, mengingat pola efek yang konsisten, hasil placebo test yang mendukung, serta robustness check yang dilakukan.