Synthetic Control Method

Professor Song Yao
Olin Business School

Customer Analytics

Q1: The Difference of Churn Rates btw Users of New Phones and Refurbished Phones

```
Biased_Estimate = df.query("newphone==1")["churn"].mean() - \
    df.query("newphone==0")["churn"].mean()
print("Biased Estimate is", Biased_Estimate)
```

```
Biased Estimate is -0.02514906044317809
```

Q2: OLS benchmark

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	0.0418	0.003	12.459	0.000	0.035	0.048
newphone	−0.0251	0.004	−6.386	0.000	−0.033	−0.017

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	0.0660	0.009	7.652	0.000	0.049	0.083
newphone	−0.0202	0.004	−4.754	0.000	−0.028	−0.012
revenue	−0.0002	8.47e−05	−2.910	0.004	−0.000	−8.05e−05
mou	−4.041e−06	5.63e−06	−0.718	0.473	−1.51e−05	7e−06
overage	0.0001	3.11e−05	3.234	0.001	3.96e−05	0.000
roam	0.0003	0.000	1.986	0.047	3.68e−06	0.001
threeway	0.0009	0.002	0.510	0.610	−0.003	0.004
months	7.312e−05	0.000	0.347	0.729	−0.000	0.000
uniqsubs	0.0018	0.002	0.875	0.381	−0.002	0.006
phones	−0.0031	0.002	−1.978	0.048	−0.006	−2.71e−05
retcalls	0.0182	0.008	2.277	0.023	0.003	0.034
dropvce	0.0001	0.000	0.462	0.644	−0.000	0.001
webcap	−0.0123	0.006	−1.993	0.046	−0.024	−0.000
children	0.0048	0.004	1.150	0.250	−0.003	0.013
mcycle	0.0014	0.016	0.085	0.932	−0.030	0.033
occcler	0.0196	0.013	1.533	0.125	−0.005	0.045
referred	−0.0166	0.004	−4.373	0.000	−0.024	−0.009

Q3: Balance Check of Features

	Variable	Treated Mean	Control Mean	Standardized Diff	p-value
0	revenue	58.452	59.545	−0.024	0.363
1	mou	527.634	519.018	0.017	0.511
2	overage	37.406	45.833	−0.083	0.002
3	roam	1.217	1.260	−0.003	0.899
4	threeway	0.305	0.241	0.068	0.003
5	months	18.222	18.602	−0.041	0.102
6	uniqsubs	1.497	1.600	−0.117	0.000
7	phones	1.717	2.361	−0.488	0.000
8	retcalls	0.023	0.093	−0.270	0.000
9	dropvce	5.901	6.556	−0.069	0.009
10	webcap	0.920	0.841	0.246	0.000
11	children	0.238	0.248	−0.023	0.384
12	mcycle	0.015	0.005	0.100	0.000
13	occcler	0.018	0.024	−0.044	0.100
14	referred	0.422	0.198	0.501	0.000

What can we conclude?

Q4: KNN

```
from causalinference import CausalModel

CM_matching = CausalModel(
    Y=df["churn"].values,
    D=df["newphone"].values,
    X=df[user_covariates].values,
)

CM_matching.est_via_matching(matches=5, bias_adj=True)

print(CM_matching.estimates)
```

```
Treatment Effect Estimates: Matching

                     Est.       S.e.          z       P>|z|      [95% Conf. int.]
--------------------------------------------------------------------------------
           ATE     -0.029      0.007     -4.089       0.000     -0.043     -0.015
           ATC     -0.023      0.006     -3.812       0.000     -0.035     -0.011
           ATT     -0.031      0.008     -3.704       0.000     -0.048     -0.015
```

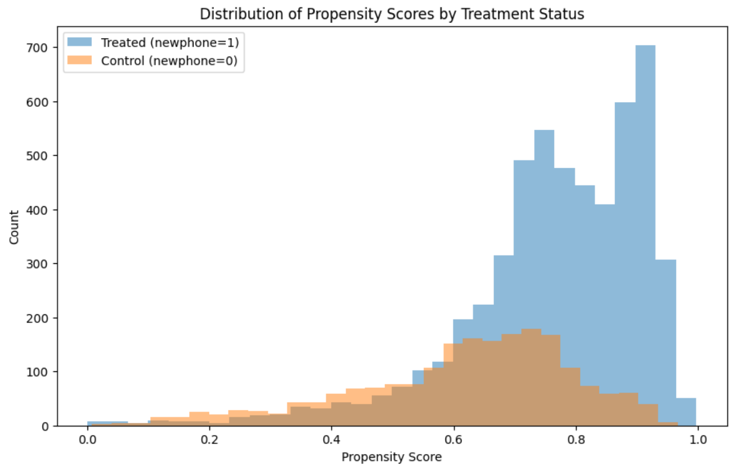
The bonus question (Q8): Which effect shall we focus on?

Q5: PSM

```
### Propensity Score Matching
# Calculate ATE using propensity score matching
cm = CausalModel(
    Y=df["churn"].values,
    D=df["newphone"].values,
    X=df["propensity_score"].values
)

cm.est_via_matching(matches=5, bias_adj=True)

print(cm.estimates)
```

```
Treatment Effect Estimates: Matching

                     Est.       S.e.          z       P>|z|      [95% Conf. int.]
--------------------------------------------------------------------------------
           ATE     -0.032      0.007     -4.276       0.000     -0.046     -0.017
           ATC     -0.016      0.006     -2.714       0.007     -0.028     -0.005
           ATT     -0.038      0.009     -4.224       0.000     -0.055     -0.020
```

Q6: SMD




```python
# Calculate standardized difference in propensity score means
treated_ps_mean = df[df['newphone']==1]['propensity_score'].mean()
control_ps_mean = df[df['newphone']==0]['propensity_score'].mean()
treated_ps_var = df[df['newphone']==1]['propensity_score'].var()
control_ps_var = df[df['newphone']==0]['propensity_score'].var()

std_diff = (treated_ps_mean - control_ps_mean) / np.sqrt((treated_ps_var + control_ps_var)/2)
print(f"\nStandardized difference in propensity score means: {std_diff:.3f}")
```

```
Standardized difference in propensity score means: 0.900
```


Q7: IPW

```
ATE using IPW:

Treatment Effect Estimates: Weighting

                     Est.       S.e.          z      P>|z|      [95% Conf. int.]
--------------------------------------------------------------------------------
           ATE     -0.023      0.012     -1.974      0.048     -0.047     -0.000
```

```
ATT: -0.0250
ATC: -0.0161
```

The bonus question (Q8): Which effect shall we focus on?