In [1]:
```
import dame_flame
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import heapq
from sklearn.model_selection import train_test_split
import statsmodels.api as sm
import statsmodels.formula.api as smf
from sklearn.preprocessing import LabelBinarizer
from sklearn.linear_model import LinearRegression
from scipy import stats
from statsmodels.distributions.empirical_distribution import ECDF
```

```
C:\Users\Neha\Anaconda3\lib\site-packages\statsmodels\tools\_testing.py:19: F
utureWarning: pandas.util.testing is deprecated. Use the functions in the pub
lic API at pandas.testing instead.
  import pandas.util.testing as tm
```

In [2]:
```
STAR_Students = pd.read_spss('STAR_Students.sav')
```

In [3]:
```
#STAR_Students['gksurban'].value_counts()
# d = {"RURAL": 0, "URBAN":1, "SUBURBAN": 2, "INNER CITY": 3}

df_trunc = STAR_Students.loc[:, STAR_Students.columns.intersection(
    ['gkclasstype', 'gender', 'race', 'gkfreelunch', 'gkschid', 'gktmathss',
'gktreadss', 'g1freelunch', 'g2freelunch', 'g3freelunch',
    'gktgen', 'gktrace', 'gkthighdegree'])]


d = {"WHITE": 1, "BLACK": 0, "ASIAN": 1, "HISPANIC": 0, "OTHER": 0,
     "NATIVE AMERICAN": 0}
df_trunc['race'] = df_trunc['race'].map(d)

d = {"NON-FREE LUNCH": 0, "FREE LUNCH": 1}
df_trunc['gkfreelunch'] = df_trunc['gkfreelunch'].map(d)
df_trunc['g1freelunch'] = df_trunc['g1freelunch'].map(d)
df_trunc['g2freelunch'] = df_trunc['g2freelunch'].map(d)
df_trunc['g3freelunch'] = df_trunc['g3freelunch'].map(d)

d = {"BACHELORS": 0, "MASTERS": 1, "MASTERS + ": 1, "SPECIALIST": 1}
df_trunc['gkthighdegree'] = df_trunc['gkthighdegree'].map(d)

d = {"MALE": 1, "FEMALE": 0}
df_trunc['gender'] = df_trunc['gender'].map(d)
df_trunc['gktgen'] = df_trunc['gktgen'].map(d)

d = {"WHITE": 1, "BLACK": 0}
df_trunc['gktrace'] = df_trunc['gktrace'].map(d)

d = {"SMALL CLASS": int(1), "REGULAR CLASS": int(0),
     "REGULAR + AIDE CLASS": int(0)}
df_trunc['ksmall'] = df_trunc['gkclasstype'].map(d)

# df_trunc = df_trunc.dropna().copy()

# Create age variable counting months
#df_trunc['age'] = df_trunc['birthyear']*12 + df_trunc['birthmonth']
df_trunc = df_trunc.drop(columns=['gkclasstype'])
# Bin age into deciles
#df_trunc['age'] = pd.qcut(df_trunc['age'], q=10, labels=False)

df_trunc = df_trunc.rename(columns={"ksmall": "treated"}) ## NOTE TO SELF -- C
OME BACK TO WE SHOULDNT HAVE TO DO THIS
```

In [4]:
```
for i in df_trunc.index:
    if df_trunc.loc[i, 'g1freelunch'] == 1 or df_trunc.loc[i, 'g2freelunch'] =
= 1 or df_trunc.loc[i, 'g3freelunch'] == 1 or df_trunc.loc[i, 'gkfreelunch'] =
= 1:
        df_trunc.loc[i, 'gkfreelunch'] = 1
    else:
        df_trunc.loc[i, 'gkfreelunch'] = 0
df_trunc = df_trunc.drop(columns=['g1freelunch', 'g2freelunch', 'g3freelunc
h'])

# df_trunc=df_trunc.dropna(subset=['treated'])
df_trunc = df_trunc.dropna() # COME BACK TO!
```

In [5]:
```
ecdf_reading = ECDF(df_trunc[df_trunc['treated'] == 0]['gktreadss'])
ecdf_math = ECDF(df_trunc[df_trunc['treated'] == 0]['gktmathss'])
df_trunc['read_outcome'] = ecdf_reading(df_trunc['gktreadss'])*100
df_trunc['math_outcome'] = ecdf_math(df_trunc['gktmathss'])*100
df_trunc['outcome'] = (df_trunc['read_outcome'] + df_trunc['math_outcome'])/2
```

In [6]:
```
'''
math_array = df_trunc[df_trunc['treated'] == 0]['gktmathss']
reading_array = df_trunc[df_trunc['treated'] == 0]['gktreadss']
#df_trunc['read_outcome'] = stats.percentileofscore(reading_array, df_trunc['g
ktreadss'])
#df_trunc['math_outcome'] = stats.percentileofscore(math_array, df_trunc['gktm
athss'])
for i in df_trunc.index:
    df_trunc.loc[i, 'math_outcome'] = stats.percentileofscore(a=math_array, sc
ore=df_trunc.loc[i,'gktmathss'])
    df_trunc.loc[i, 'read_outcome'] = stats.percentileofscore(a=reading_array,
score=df_trunc.loc[i,'gktreadss'])


df_trunc['outcome'] = (df_trunc['read_outcome'] + df_trunc['math_outcome'])/2
'''

'''
df_trunc_untreated = df_trunc[df_trunc['treated'] == 0]
df_trunc_treated = df_trunc[df_trunc['treated'] == 1]
for i in df_trunc_treated.index:
    df_trunc_treated.loc[i, 'gktreadss'] = stats.percentileofscore(df_trunc_un
treated['gktreadss'], df_trunc_treated.loc[i,'gktreadss'])
    df_trunc_treated.loc[i, 'gktmathss'] = stats.percentileofscore(df_trunc_un
treated['gktmathss'], df_trunc_treated.loc[i,'gktmathss'])

# Percentile the math and reading and then average them
# but do different percentiles for the small class size people and the large p
eople.
df_trunc_untreated = df_trunc[df_trunc['treated'] == 0]
df_trunc_untreated['gktreadss'] = df_trunc_untreated['gktreadss'].rank(pct=Tru
e)*100
df_trunc_untreated['gktmathss'] = df_trunc_untreated['gktmathss'].rank(pct=Tru
e)*100
df_trunc_untreated['outcome'] = df_trunc_untreated[['gktreadss', 'gktmaths
s']].mean(axis=1)

#df_trunc_treated = df_trunc[df_trunc['treated'] == 1]
#df_trunc_treated['gktreadss'] = df_trunc_treated['gktreadss'].rank(pct=True)*
100
#df_trunc_treated['gktmathss'] = df_trunc_treated['gktmathss'].rank(pct=True)*
100
df_trunc_treated['outcome'] = df_trunc_treated[['gktreadss', 'gktmathss']].mea
n(axis=1)

df = pd.concat([df_trunc_treated, df_trunc_untreated])
df = df.drop(columns=['gktreadss', 'gktmathss'])

df = df.dropna()
'''
```

Out[6]:
```
"\ndf_trunc_untreated = df_trunc[df_trunc['treated'] == 0]\ndf_trunc_treated
= df_trunc[df_trunc['treated'] == 1]\nfor i in df_trunc_treated.index:\n    d
f_trunc_treated.loc[i, 'gktreadss'] = stats.percentileofscore(df_trunc_untrea
ted['gktreadss'], df_trunc_treated.loc[i,'gktreadss'])\n    df_trunc_treated.
loc[i, 'gktmathss'] = stats.percentileofscore(df_trunc_untreated['gktmaths
s'], df_trunc_treated.loc[i,'gktmathss'])\n\n# Percentile the math and readin
g and then average them\n# but do different percentiles for the small class s
ize people and the large people.\ndf_trunc_untreated = df_trunc[df_trunc['tre
ated'] == 0]\ndf_trunc_untreated['gktreadss'] = df_trunc_untreated['gktreads
s'].rank(pct=True)*100\ndf_trunc_untreated['gktmathss'] = df_trunc_untreated
['gktmathss'].rank(pct=True)*100\ndf_trunc_untreated['outcome'] = df_trunc_un
treated[['gktreadss', 'gktmathss']].mean(axis=1)\n\n#df_trunc_treated = df_tr
unc[df_trunc['treated'] == 1]\n#df_trunc_treated['gktreadss'] = df_trunc_trea
ted['gktreadss'].rank(pct=True)*100\n#df_trunc_treated['gktmathss'] = df_trun
c_treated['gktmathss'].rank(pct=True)*100\ndf_trunc_treated['outcome'] = df_t
runc_treated[['gktreadss', 'gktmathss']].mean(axis=1)\n\ndf = pd.concat([df_t
runc_treated, df_trunc_untreated])\ndf = df.drop(columns=['gktreadss', 'gktma
thss'])\n\ndf = df.dropna()\n"
```

In [7]:
```
md = sm.OLS(df_trunc['outcome'], sm.add_constant(df_trunc['treated']))
md.fit().summary()
```

Out[7]:

OLS Regression Results

Dep. Variable:	outcome	**R-squared:**	0.006
Model:	OLS	**Adj. R-squared:**	0.006
Method:	Least Squares	**F-statistic:**	36.27
Date:	Mon, 13 Feb 2023	**Prob (F-statistic):**	1.83e-09
Time:	22:30:00	**Log-Likelihood:**	-26366.
No. Observations:	5593	**AIC:**	5.274e+04
Df Residuals:	5591	**BIC:**	5.275e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	51.3446	0.432	118.717	0.000	50.497	52.192
treated	4.7260	0.785	6.023	0.000	3.188	6.264

Omnibus:	2138.187	**Durbin-Watson:**	1.801
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	288.714
Skew:	-0.113	**Prob(JB):**	2.03e-63
Kurtosis:	1.910	**Cond. No.**	2.42

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [8]:
```
md = smf.mixedlm(formula="outcome ~ treated", data=df_trunc , groups=df_trunc
['gkschid'])
md.fit().summary()
```

Out[8]:

Model:	MixedLM	Dependent Variable:	outcome
No. Observations:	5593	Method:	REML
No. Groups:	79	Scale:	557.4868
Min. group size:	34	Likelihood:	-25738.8904
Max. group size:	134	Converged:	Yes
Mean group size:	70.8		

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	51.090	1.532	33.355	0.000	48.088	54.093
treated	5.301	0.700	7.574	0.000	3.929	6.673
Group Var	173.240	1.242				

In [9]:
```
md = smf.mixedlm(formula="outcome ~ gender+race+gkfreelunch+treated", data=df_
trunc, groups=df_trunc['gkschid'])
md.fit().summary()
```

Out[9]:

Model:	MixedLM	Dependent Variable:	outcome
No. Observations:	5593	Method:	REML
No. Groups:	79	Scale:	514.9750
Min. group size:	34	Likelihood:	-25510.8027
Max. group size:	134	Converged:	Yes
Mean group size:	70.8		

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	41.477	1.688	24.566	0.000	38.167	44.786
gender[T.1]	-4.835	0.611	-7.911	0.000	-6.033	-3.637
gkfreelunch[T.0]	12.419	0.735	16.893	0.000	10.978	13.860
race	9.314	1.230	7.571	0.000	6.903	11.726
treated	5.361	0.673	7.970	0.000	4.043	6.680
Group Var	148.774	1.116				

In [10]:
```
fes = pd.get_dummies(df_trunc['gkschid'])
fes = fes.drop(columns=[161183.0])
y = df_trunc.loc[:,['outcome']]
x = df_trunc.loc[:, ['gender', 'race', 'gkfreelunch', 'treated']]
x = pd.concat([fes,x],axis=1)
x = sm.add_constant(x)
model = sm.OLS(y,x)
model.fit().summary()
```

Out[10]:

OLS Regression Results

Dep. Variable:	outcome	**R-squared:**	0.308
Model:	OLS	**Adj. R-squared:**	0.297
Method:	Least Squares	**F-statistic:**	29.87
Date:	Mon, 13 Feb 2023	**Prob (F-statistic):**	0.00
Time:	22:30:02	**Log-Likelihood:**	-25356.
No. Observations:	5593	**AIC:**	5.088e+04
Df Residuals:	5510	**BIC:**	5.143e+04
Df Model:	82		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	61.9870	2.640	23.479	0.000	56.811	67.163
112038.0	-31.3366	3.871	-8.096	0.000	-38.925	-23.748
123056.0	-18.0902	3.863	-4.683	0.000	-25.662	-10.518
128068.0	-19.9066	3.789	-5.253	0.000	-27.335	-12.478
128076.0	-27.0785	3.711	-7.297	0.000	-34.354	-19.803
128079.0	-27.7731	3.683	-7.541	0.000	-34.993	-20.553
130085.0	-18.6943	3.493	-5.351	0.000	-25.543	-11.846
159171.0	8.1143	3.206	2.531	0.011	1.828	14.400
161176.0	-17.9768	3.385	-5.311	0.000	-24.612	-11.341
162184.0	-14.2118	3.664	-3.879	0.000	-21.395	-7.029
164198.0	-9.0783	3.792	-2.394	0.017	-16.513	-1.644
165199.0	0.7355	4.257	0.173	0.863	-7.611	9.082
166203.0	-19.7471	4.011	-4.923	0.000	-27.611	-11.883
168211.0	-10.2904	3.209	-3.207	0.001	-16.580	-4.000
168214.0	2.0569	3.773	0.545	0.586	-5.340	9.454
169219.0	1.1050	3.963	0.279	0.780	-6.665	8.875
169229.0	-3.8031	3.257	-1.168	0.243	-10.189	2.583
169231.0	-28.2731	3.848	-7.348	0.000	-35.816	-20.730
169280.0	-4.7299	3.828	-1.236	0.217	-12.233	2.774
170295.0	-4.3749	3.588	-1.219	0.223	-11.410	2.660
173312.0	12.2725	3.750	3.272	0.001	4.920	19.625
176329.0	3.7140	3.556	1.044	0.296	-3.258	10.686
180344.0	-8.5882	3.259	-2.636	0.008	-14.976	-2.200
189378.0	-23.7440	3.514	-6.757	0.000	-30.632	-16.856
189382.0	-11.6621	3.623	-3.219	0.001	-18.765	-4.560

189396.0	-22.3182	3.661	-6.096	0.000	-29.495	-15.141
191411.0	-4.6793	4.068	-1.150	0.250	-12.654	3.296
193422.0	0.9670	3.726	0.260	0.795	-6.337	8.271
193423.0	-4.5623	3.462	-1.318	0.188	-11.349	2.224
201449.0	3.9790	3.252	1.224	0.221	-2.395	10.353
203452.0	-14.4261	3.259	-4.426	0.000	-20.816	-8.036
203457.0	3.2697	4.170	0.784	0.433	-4.905	11.444
205488.0	-12.5168	3.822	-3.275	0.001	-20.010	-5.024
205489.0	-13.2046	3.886	-3.398	0.001	-20.822	-5.587
205490.0	-29.9229	3.866	-7.739	0.000	-37.502	-22.343
205491.0	-17.1038	3.554	-4.813	0.000	-24.071	-10.137
205492.0	9.0350	3.576	2.526	0.012	2.024	16.046
208501.0	-12.6391	3.629	-3.482	0.001	-19.754	-5.524
208503.0	-29.5522	4.037	-7.321	0.000	-37.466	-21.638
209510.0	-19.6675	3.249	-6.054	0.000	-26.037	-13.298
212522.0	-2.8397	3.588	-0.791	0.429	-9.874	4.195
215533.0	3.3418	3.130	1.068	0.286	-2.794	9.477
216536.0	-14.3392	3.182	-4.506	0.000	-20.578	-8.101
218562.0	-2.1619	3.751	-0.576	0.564	-9.516	5.192
221571.0	-39.4511	3.215	-12.270	0.000	-45.754	-33.148
221574.0	-25.9991	3.659	-7.105	0.000	-33.173	-18.825
225585.0	-17.9102	3.793	-4.722	0.000	-25.346	-10.475
228606.0	-6.8059	3.434	-1.982	0.048	-13.538	-0.074
230612.0	7.7796	3.702	2.101	0.036	0.522	15.038
231616.0	-0.5930	3.744	-0.158	0.874	-7.932	6.746
234628.0	-6.8863	3.204	-2.149	0.032	-13.168	-0.605
244697.0	-17.3363	3.432	-5.051	0.000	-24.065	-10.608
244708.0	-21.9966	3.349	-6.568	0.000	-28.562	-15.432
244723.0	-19.2621	3.568	-5.399	0.000	-26.257	-12.268
244727.0	-3.9339	3.608	-1.090	0.276	-11.008	3.140
244728.0	-17.2271	4.108	-4.194	0.000	-25.280	-9.174
244736.0	12.2747	4.081	3.008	0.003	4.274	20.276
244745.0	6.1293	3.450	1.777	0.076	-0.634	12.893
244746.0	2.3925	4.033	0.593	0.553	-5.513	10.298
244755.0	1.1748	3.217	0.365	0.715	-5.131	7.481
244764.0	-6.1613	4.639	-1.328	0.184	-15.255	2.932
244774.0	-3.7358	3.401	-1.099	0.272	-10.402	2.931

244776.0	-8.1902	3.211	-2.551	0.011	-14.485	-1.896
244780.0	30.6345	3.903	7.848	0.000	22.982	38.287
244796.0	-10.5236	3.968	-2.652	0.008	-18.301	-2.746
244799.0	-10.7156	3.819	-2.806	0.005	-18.203	-3.228
244801.0	-12.4134	3.523	-3.524	0.000	-19.319	-5.508
244806.0	19.1338	3.194	5.991	0.000	12.873	25.395
244818.0	-17.0011	3.504	-4.852	0.000	-23.870	-10.133
244831.0	-10.3377	3.802	-2.719	0.007	-17.791	-2.885
244839.0	9.4845	3.630	2.613	0.009	2.368	16.601
252885.0	-4.8898	3.564	-1.372	0.170	-11.877	2.097
253888.0	-11.0419	4.247	-2.600	0.009	-19.367	-2.717
257899.0	-18.4010	3.161	-5.821	0.000	-24.598	-12.204
257905.0	-2.3882	3.023	-0.790	0.430	-8.314	3.538
259915.0	-13.3461	3.723	-3.585	0.000	-20.645	-6.048
261927.0	-8.9742	3.344	-2.684	0.007	-15.529	-2.420
262937.0	5.2583	3.584	1.467	0.142	-1.768	12.284
264945.0	-0.8040	3.233	-0.249	0.804	-7.141	5.533
gender	-4.7859	0.611	-7.828	0.000	-5.984	-3.587
race	9.8631	1.298	7.598	0.000	7.318	12.408
gkfreelunch	-12.4237	0.739	-16.808	0.000	-13.873	-10.975
treated	5.3851	0.673	7.997	0.000	4.065	6.705

Omnibus:	122.978	**Durbin-Watson:**	1.883
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	80.379
Skew:	-0.166	**Prob(JB):**	3.52e-18
Kurtosis:	2.515	**Cond. No.**	98.5

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

In [11]:
```
df_trunc = df_trunc.drop(columns=['math_outcome', 'read_outcome'])
```

In [12]:
```
# Do mice
'''
from sklearn.tree import DecisionTreeRegressor

from sklearn.experimental import enable_iterative_imputer
from sklearn.impute import IterativeImputer

imp = IterativeImputer(max_iter=10, random_state=1,
                       estimator=DecisionTreeRegressor())
imp.fit(df)
tmp_df = pd.DataFrame(data=np.round(imp.transform(df)),
                      columns=df.columns,
                      index=df.index)
# convert floats to ints because MICE creates floats
cols = list(df.columns)
cols.remove("outcome")
tmp_df[cols] = tmp_df[cols].astype('int64')
'''
```

Out[12]:
```
'\nfrom sklearn.tree import DecisionTreeRegressor\n\nfrom sklearn.experimenta
l import enable_iterative_imputer\nfrom sklearn.impute import IterativeImpute
r\n\nimp = IterativeImputer(max_iter=10, random_state=1,\n
estimator=DecisionTreeRegressor())\nimp.fit(df)\ntmp_df = pd.DataFrame(data=n
p.round(imp.transform(df)),\n                      columns=df.columns,\n
index=df.index)\n# convert floats to ints because MICE creates floats\ncols =
list(df.columns)\ncols.remove("outcome")\ntmp_df[cols] = tmp_df[cols].astype
(\'int64\')\n'
```

In [13]:
```
df_trunc = df_trunc.drop(columns=['gktreadss', 'gktmathss'])
```

In [14]:
```
df_trunc.head()
```

Out[14]:

	gender	race	gkschid	gktgen	gktrace	gkthighdegree	gkfreelunch	treated	outcome
133	1	0.0	169280.0	0	1	0.0	0	0.0	48.292244
246	0	1.0	218562.0	0	1	1.0	1	0.0	70.107858
263	0	0.0	205492.0	0	1	0.0	1	1.0	85.618901
266	1	1.0	257899.0	0	1	1.0	0	0.0	79.314330
275	1	1.0	161176.0	0	1	0.0	1	0.0	32.383154

In [15]:
```
# Do the matching

models = []
random_seeds = [1111, 2222, 3333, 4444]
for i in range(4):
    matching_df, holdout_df = train_test_split(df_trunc, test_size=0.2, random
_state=random_seeds[i])
    model_dame = dame_flame.matching.DAME(
        repeats=False, verbose=0, adaptive_weights='decisiontree')
    model_dame.fit(holdout_data=holdout_df)
    model_dame.predict(matching_df)
    models.append(model_dame)
```

```
4363 units matched. We finished with no more treated units to match
4290 units matched. We finished with no more treated units to match
4376 units matched. We finished with no more treated units to match
4379 units matched. We finished with no more treated units to match
```

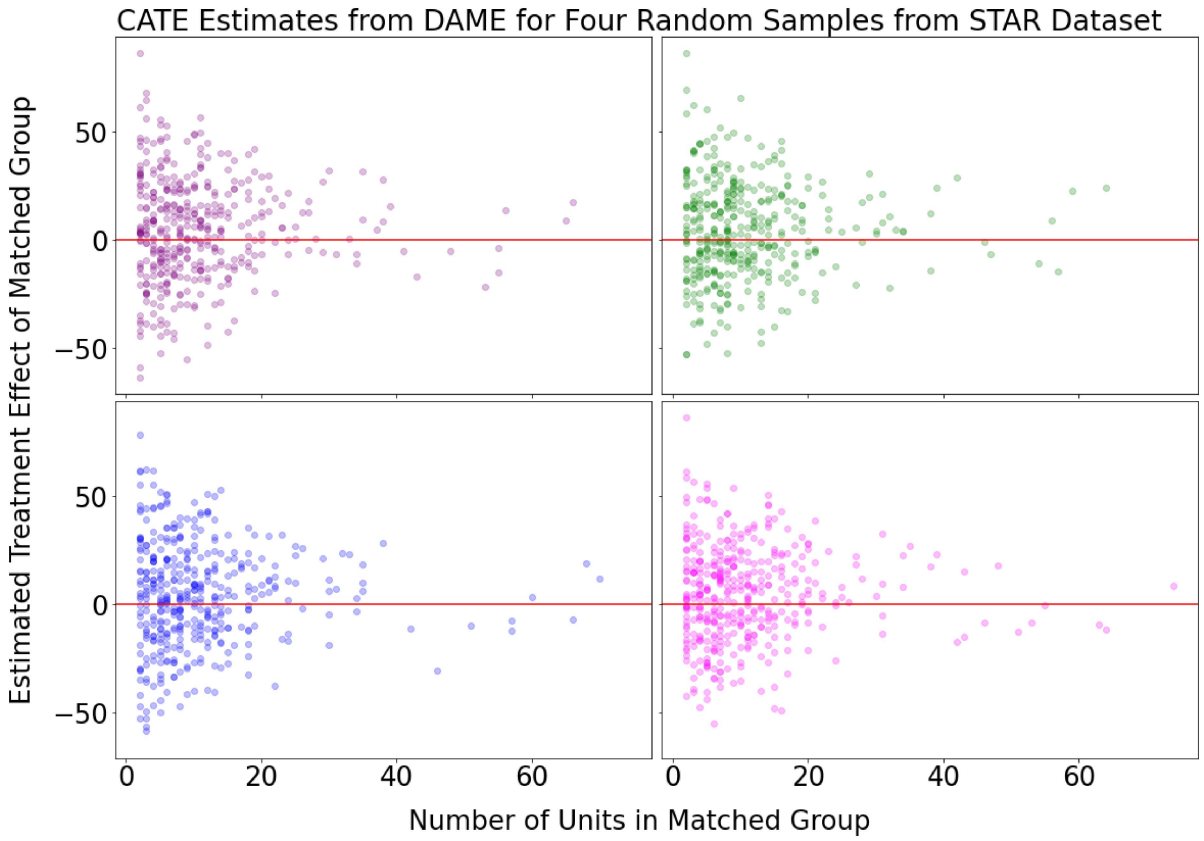
In [16]:
```
for i in range(len(models)):
    ate, var = dame_flame.utils.post_processing.var_ATE(matching_object=models
[i])
    print("ATE of trial", i, ":", ate,". Variance: ", var)
```

```
treated_col treated
ATE of trial 0 : 4.98772070404948 . Variance:  0.873569213376045
treated_col treated
ATE of trial 1 : 4.5776967755522096 . Variance:  0.8649523950028482
treated_col treated
ATE of trial 2 : 4.274075245409583 . Variance:  0.9710383214554565
treated_col treated
ATE of trial 3 : 4.653545498490466 . Variance:  0.9213514936610683
```

In [17]:
```
# compute stuff for plot
# Create the plot
match_dfs = []
for i in models:
    match_dfs.append(i.input_data)

for i in range(4):
    colname = 'cates'
    match_dfs[i][colname] = dame_flame.utils.post_processing.CATE(
        models[i], match_dfs[i].index)

dame_len_groups = []
dame_cate_of_groups = []

for i in range(4):

    model_dame = models[i]
    groups = list(range(len(model_dame.units_per_group)))

    dame_cate_of_group = []
    dame_len_group = []
    dame_len_treated = []
    maxcate = 0.0
    maxgroupnum = 0
    index = 0

    flame_cate_of_group = []
    flame_len_group = []
    large_groups = []
    for group in model_dame.units_per_group:
        dame_cate_of_group.append(dame_flame.utils.post_processing.CATE(
            model_dame, group[0]))
        dame_len_group.append(len(group))

        # find len of just treated units
        df_mmg = df_trunc.loc[group]
        treated = df_mmg.loc[df_mmg["treated"] == 1]

    dame_len_groups.append(dame_len_group)
    dame_cate_of_groups.append(dame_cate_of_group)
```

```
C:\Users\Neha\Anaconda3\lib\site-packages\ipykernel_launcher.py:10: SettingWi
thCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/s
table/user_guide/indexing.html#returning-a-view-versus-a-copy
  # Remove the CWD from sys.path while we load stuff.
C:\Users\Neha\Anaconda3\lib\site-packages\ipykernel_launcher.py:10: SettingWi
thCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/s
table/user_guide/indexing.html#returning-a-view-versus-a-copy
  # Remove the CWD from sys.path while we load stuff.
C:\Users\Neha\Anaconda3\lib\site-packages\ipykernel_launcher.py:10: SettingWi
thCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/s
table/user_guide/indexing.html#returning-a-view-versus-a-copy
  # Remove the CWD from sys.path while we load stuff.
C:\Users\Neha\Anaconda3\lib\site-packages\ipykernel_launcher.py:10: SettingWi
thCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/s
table/user_guide/indexing.html#returning-a-view-versus-a-copy
  # Remove the CWD from sys.path while we load stuff.
```

In [18]:
```
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplots(2, 2, figsize = (19,13),
                                             sharex=True, sharey=True)
fig.text(0.5, 0.05, 'Number of Units in Matched Group', ha='center',
         fontsize=26)
fig.text(0.05, 0.5, 'Estimated Treatment Effect of Matched Group',
         va='center', rotation='vertical', fontsize=26)
fig.suptitle("CATE Estimates from DAME for Four Random Samples from STAR Datas
et", fontsize=28, y=0.91)
ax1.axhline(y=0.0, color='r', linestyle='-')
ax2.axhline(y=0.0, color='r', linestyle='-')
ax3.axhline(y=0.0, color='r', linestyle='-')
ax4.axhline(y=0.0, color='r', linestyle='-')

ax1.tick_params(labelsize=26)
ax2.tick_params(labelsize=26)
ax3.tick_params(labelsize=26)
ax4.tick_params(labelsize=26)

ax1.scatter(dame_len_groups[0], dame_cate_of_groups[0], color="purple",
            alpha = 0.25)
#ax1.text(0.15, 0.9,'ATE: '+str(round(ates[0],2)), ha='center', va='center',
#         transform=ax1.transAxes, fontsize=26)

ax2.scatter(dame_len_groups[1], dame_cate_of_groups[1], color="green",
            alpha = 0.25)
#ax2.text(0.15, 0.9,'ATE: '+str(round(ates[1],2)), ha='center', va='center',
#         transform=ax2.transAxes, fontsize=26)

ax3.scatter(dame_len_groups[2], dame_cate_of_groups[2], color="blue",
            alpha = 0.25)
#ax3.text(0.15, 0.9,'ATE: '+str(round(ates[2],2)), ha='center', va='center',
#         transform=ax3.transAxes, fontsize=26)

ax4.scatter(dame_len_groups[3], dame_cate_of_groups[3], color="magenta",
            alpha = 0.25)
#ax4.text(0.15, 0.9,'ATE: '+str(round(ates[3],2)), ha='center', va='center',
#         transform=ax4.transAxes, fontsize=26)

plt.subplots_adjust(wspace=.02, hspace=.02)
## plt.savefig('cate-graph4.png', dpi = 200)
```



In [19]:
```
list_star_covars = []
for modelid in range(len(models)):

    # Pull out the groups with 10 or more units in the matched group
    model = models[modelid]
    large_groups = []
    for group in model.units_per_group:
        if len(group) >= 12.5:
            large_groups.append(group)

    covariates = set(models[modelid].input_data.columns) - set(['outcome', 'tr
eated', 'cates'])
    # Which covars did the large group match on?
    star_covars = dict()
    for group in large_groups:
        group_star_covars = []
        matched_df = models[modelid].df_units_and_covars_matched.loc[group]
        for covar in covariates:
            if '*' in matched_df[covar].values:
                group_star_covars.append(covar)
        cate_of_group = models[modelid].input_data.loc[group[0], 'cates']
        star_covars[cate_of_group] = group_star_covars

    list_star_covars.append(star_covars)
```

In [20]:
```
list_star_covars
```

Out[20]:
```
[{-12.006933744221882: [],
  5.609199338012914: [],
  40.1296866974833: [],
  -23.25515037379445: [],
  27.96197446621175: [],
  6.915224605902559: [],
  12.887776065742166: [],
  40.031458654340014: [],
  30.036028278799975: [],
  -20.286908634366256: [],
  -42.084892508621316: [],
  22.421372246895686: [],
  -2.792758089368249: [],
  13.077228369986457: [],
  13.99375107002227: [],
  27.580037664783426: [],
  -9.62477886206701: [],
  -37.309536038349606: [],
  -28.626458287475238: [],
  -21.062318096216394: [],
  -17.40223053782376: [],
  -14.043685441990519: [],
  32.7914740626605: [],
  2.0081260547362163: [],
  21.357157865247245: [],
  -0.9972607430234603: [],
  15.949122329979737: [],
  15.96580208868344: [],
  -5.251669234720083: [],
  23.62772563851147: [],
  -0.013642783769896027: [],
  21.018554185927073: [],
  0.784808411927056: [],
  -6.545682816869245: [],
  19.493451463790457: [],
  41.89082104336342: [],
  39.70810904555126: [],
  -10.444843919420194: [],
  5.45550847457627: [],
  4.5957455915083045: [],
  29.18366943790673: [],
  -12.818672082924778: [],
  -29.424756034925537: [],
  3.032871083718547: [],
  1.748110646415732: [],
  4.276094276094273: [],
  16.877461051189876: [],
  22.487801746276332: [],
  -17.97530388632083: [],
  32.30985704502653: [],
  -5.335665981852422: [],
  18.859784283513086: [],
  10.789210789210784: [],
  37.1003466872111: [],
  27.149960052502436: [],
  31.78222907036468: [],
  10.647149460708782: [],
```

```
 0.18538135593220773: [],
 8.15838469440164: [],
 23.760330578512395: [],
 31.795629041391763: [],
 -11.110041088854643: [],
 -10.185541859270671: [],
 -7.92664256647307: [],
 -24.259544598527647: [],
 26.550248245163502: [],
 6.097292537970503: [],
 18.193129845672217: [],
 17.950693374422194: [],
 -12.336439455083521: [],
 6.025868552351611: [],
 12.7788352115371: [],
 15.758790997138448: [],
 -4.8947098099640485: ['gkthighdegree'],
 15.72019631341665: ['gkthighdegree'],
 0.047651657821162985: ['gkthighdegree'],
 -13.581500677032267: ['gkthighdegree'],
 10.506301607996527: ['gkthighdegree'],
 13.347691086520065: ['gkthighdegree'],
 -4.87373737373737: ['gkthighdegree'],
 -3.0192137087360464: ['gkthighdegree'],
 -8.62758945386063: ['gkthighdegree'],
 -6.9658448895737095: ['gkthighdegree'],
 12.162871654397073: ['gkthighdegree'],
 9.458037416833452: ['gkthighdegree'],
 8.425345385432877: ['gkthighdegree'],
 -12.323446327683598: ['gkthighdegree'],
 13.388475146949723: ['gkthighdegree'],
 0.6549652754516302: ['gkthighdegree'],
 12.773314256365097: ['gkthighdegree'],
 23.542629686697484: ['gktrace'],
 21.777721508075896: ['gktrace'],
 11.317472546286105: ['gktrace'],
 -15.117656590114237: ['gktrace'],
 -16.631465366297803: ['gktrace'],
 1.4481986939614018: ['gkthighdegree', 'gktrace'],
 0.476257178876601: ['gkschid'],
 -10.806154768019162: ['gkschid'],
 -6.292953750580871: ['gkschid'],
 28.018889459567433: ['gkschid'],
 -6.296622398317314: ['gkschid'],
 -3.6861196712891555: ['gkschid'],
 5.080660223187181: ['gkschid'],
 -4.895491391438327: ['gkschid'],
 -21.687246761399308: ['gkschid'],
 -3.7934993443467846: ['gkschid'],
 -6.491338656207688: ['gkschid'],
 -4.5005136106831145: ['gkschid'],
 17.662643382982367: ['gkschid'],
 9.26425269645609: ['gkschid'],
 13.73694572847117: ['gkschid'],
 -4.899601340279297: ['gkschid']},
{18.23898794237776: [],
 -19.492327940421184: [],
```

```
10.104220167779488: [],
18.214836011446195: [],
10.033812703304221: [],
3.1408719968041936: [],
29.95695453322572: [],
7.949730354391356: [],
-3.2978085944187754: [],
37.520330422872796: [],
-30.67682474462135: [],
-39.98826032724338: [],
-0.4245848313644984: [],
2.6397449066940766: [],
25.65855006147764: [],
-24.93223192375737: [],
-14.438345317582602: [],
-32.244764024425045: [],
-32.964175654853626: [],
-1.4682845403184501: [],
-23.69029275808937: [],
-17.360897106659838: [],
-7.774781715459682: [],
25.151698583901982: [],
-1.3208190114969796: [],
20.77662215374079: [],
-3.867254984358226: [],
9.348570450265342: [],
23.35965588084233: [],
-4.772299263824685: [],
23.777840033618162: [],
-5.823519700638364: [],
21.35036808765623: [],
4.196185693360828: [],
-24.25954459852765: [],
20.01551208022702: [],
45.877807349301946: [],
41.75333846944017: [],
4.318609827084394: [],
-11.436825885978436: [],
3.9307522358369837: [],
9.187794742494283: [],
2.6896079438452176: [],
29.182564140191275: [],
-47.29420553765701: [],
-42.14058925152918: [],
-14.668578439764893: [],
-10.828625235404907: [],
10.594443221561868: [],
43.74831853645413: [],
8.113336757404554: [],
20.476957557080837: [],
-16.935534077290626: [],
21.50082967879578: [],
4.327170005136104: [],
14.164673046251998: [],
16.395798175459184: [],
35.2877035080425: [],
36.24539890429721: [],
```

```
30.57200888479779: [],
24.052783374817288: [],
11.100410888546485: [],
-11.964275523597564: [],
-20.714502498015595: [],
39.28041976000374: [],
26.87949409347715: [],
-19.004736631855266: [],
-7.476926429160638: [],
-14.856435936944408: [],
-27.54604152909237: [],
20.173129601095702: [],
-2.062224712455837: [],
5.700261474529569: [],
17.714065595421516: [],
34.069508645779834: [],
16.32055013410946: [],
-14.857717611954897: [],
4.183457785152697: [],
5.642013353877772: [],
12.245655709638768: [],
24.159765023112477: ['gktrace'],
20.850259139935574: ['gktrace'],
-10.277716633648833: ['gktrace'],
19.161958568738235: ['gktrace'],
18.087350502604735: ['gktrace'],
18.238957370313308: ['gktrace'],
-10.464817668207502: ['gktrace'],
-13.858591972998752: ['gktrace'],
-3.2428094504365745: ['gkthighdegree'],
-5.571962563487986: ['gkthighdegree'],
25.571086164306507: ['gkthighdegree'],
-5.29789419619928: ['gkthighdegree'],
-17.332526230831316: ['gkthighdegree'],
11.87527162103433: ['gkthighdegree'],
-6.341610367034086: ['gkthighdegree'],
-2.882639958911156: ['gkthighdegree'],
18.054057524396505: ['gkthighdegree'],
7.083241617139919: ['gkthighdegree'],
-3.1811761684643045: ['gkthighdegree'],
-2.0833333333333286: ['gkthighdegree'],
6.081150487930145: ['gkthighdegree'],
13.622025338127038: ['gkthighdegree'],
4.85288722576858: ['gkthighdegree'],
8.215459681561374: ['gkthighdegree'],
-0.8684486724280447: ['gkthighdegree'],
4.729070364663592: ['gkthighdegree'],
-27.30366064341412: ['gkthighdegree', 'gktrace'],
-14.971751412429377: ['gkschid'],
4.004126596178018: ['gkschid'],
10.891297967569152: ['gkschid'],
29.12767609956532: ['gkschid'],
-21.831450094161962: ['gkschid'],
-1.212793797539561: ['gkschid'],
-1.0289723117149947: ['gkschid'],
-14.58235247959827: ['gkschid'],
-6.494909211758674: ['gkschid'],
```

```
 -4.876733436055481: ['gkschid'],
 4.092835130970727: ['gkschid'],
 8.903783598698865: ['gkschid'],
 24.03634125950022: ['gkschid'],
 22.807681752721656: ['gkschid']},
{-10.77683615819209: [],
 5.599867065470271: [],
 52.75873138161274: [],
 19.31327230429921: [],
 23.838872991415357: [],
 3.471152199965772: [],
 7.468426112493894: [],
 22.99750198440492: [],
 35.56753980482796: [],
 22.717000513610685: [],
 -5.379358557324657: [],
 9.700465103007488: [],
 39.3969354562575: [],
 -28.029982023626086: [],
 -40.4275808936826: [],
 -18.476341807909613: [],
 4.742950081933131: [],
 16.675421394219555: [],
 23.66782229070364: [],
 -11.299435028248588: [],
 -18.048065399760326: [],
 -27.23848656052047: [],
 -24.056240369799696: [],
 -12.192917508788064: [],
 24.168847473932217: [],
 -3.8209522030785763: [],
 21.144709809964056: [],
 -2.3811561947155013: [],
 17.95871854134566: [],
 14.721201585608362: [],
 7.275163631095829: [],
 31.251070022256464: [],
 8.507135519847388: [],
 20.8083661473492: [],
 2.735282608163967: [],
 -11.936862572455787: [],
 10.95388917422816: [],
 1.754622496147924: [],
 -8.437889793821995: [],
 4.050187101034581: [],
 26.88364829753165: [],
 3.193374422187979: [],
 -14.02028762198254: [],
 -22.03389830508474: [],
 16.76321559796137: [],
 4.00391628145865: [],
 17.44649888717685: [],
 -29.19555726759117: [],
 7.854819380243105: [],
 23.534727983880526: [],
 7.119997878564568: [],
 -12.157592877931862: [],
```

```
12.92408548764481: [],
11.09934086629002: [],
32.27148910411623: [],
26.183138895003296: [],
18.299715179530274: [],
5.983366915570301: [],
-13.571978868589042: [],
-16.13222621697198: [],
32.119010932570255: [],
32.096655110426305: [],
-10.986988529361405: [],
-9.81448732315451: [],
-10.441888619854709: [],
-32.387104421002725: [],
9.054654219160689: [],
9.768514764054473: [],
21.324549979210985: [],
19.08577298407807: [],
-18.846944016435558: [],
2.1780303030303116: [],
11.49052841475573: [],
10.022049837946994: [],
21.289162814586554: ['gktrace'],
31.638061405010564: ['gktrace'],
15.57982977474503: ['gktrace'],
13.276836158192097: ['gktrace'],
-7.639439836518839: ['gktrace'],
-0.16649546310562613: ['gkthighdegree'],
-2.8961935741596747: ['gkthighdegree'],
29.62827426810479: ['gkthighdegree'],
24.489307559415415: ['gkthighdegree'],
2.815870570107869: ['gkthighdegree'],
-17.368212167904: ['gkthighdegree'],
50.1070022256463: ['gkthighdegree'],
-2.2119500085601764: ['gkthighdegree'],
17.165102488677213: ['gkthighdegree'],
18.331086286594775: ['gkthighdegree'],
-11.651258346173606: ['gkthighdegree'],
-7.3790680300695755: ['gkthighdegree'],
-0.37741213588672906: ['gkthighdegree'],
4.363122752953274: ['gkthighdegree'],
6.869251062240281: ['gkthighdegree'],
11.471494607087827: ['gkthighdegree'],
6.308113277721695: ['gkthighdegree'],
-18.821453708459337: ['gkthighdegree'],
8.46976117103236: ['gkthighdegree'],
-4.66476202705017: ['gkthighdegree', 'gktrace'],
-14.409697903534564: ['gender', 'gkthighdegree'],
-1.6530448181148287: ['gkschid'],
-3.4334095967378033: ['gkschid'],
-7.093729807644529: ['gkschid'],
-1.0496918335901384: ['gkschid'],
28.23503395537295: ['gkschid'],
-37.74793944285471: ['gkschid'],
-30.659704388517937: ['gkschid'],
-9.926574636017733: ['gkschid'],
-12.18932136397136: ['gkschid'],
```

```
 -11.352352492568215: ['gkschid'],
 6.040156746371437: ['gkschid'],
 -16.856479310421832: ['gkschid'],
 11.71099233864068: ['gkschid'],
 3.486353895471346: ['gkschid'],
 19.05377931860984: ['gkschid'],
 -4.589111453518228: ['gkschid']},
{-11.808376523322593: [],
 5.330070394875975: [],
 45.935627461051205: [],
 50.74673286537694: [],
 22.189815833883628: [],
 13.289676425269647: [],
 0.07645795396179977: [],
 9.661261525668309: [],
 28.344889573703135: [],
 38.63183177739509: [],
 -2.819722650231121: [],
 -10.405813583779697: [],
 12.829061026287519: [],
 39.8733093648348: [],
 -47.888693227676285: [],
 -49.11744564286938: [],
 6.554344901802523: [],
 15.092770929635364: [],
 35.047294983735654: [],
 -7.285934404578491: [],
 -40.23658192090397: [],
 -37.99526744441998: [],
 -32.06892370027964: [],
 -21.116313065465604: [],
 -8.477144324601944: [],
 27.45409604519775: [],
 -6.977722136620443: [],
 24.165382639958906: [],
 -2.7446070878274043: [],
 11.356788221194996: [],
 22.82571477486733: [],
 -9.2463263495621: [],
 18.778594287068884: [],
 0.8360989293192631: [],
 19.746083718541342: [],
 1.2056399344534903: [],
 -10.498345032243343: [],
 18.484439930896038: [],
 42.90295391990307: [],
 36.068310220852595: [],
 9.25977805792904: [],
 -15.182515224888093: [],
 10.059765606760987: [],
 2.43323061119672: [],
 25.669120584374824: [],
 -18.650682479650122: [],
 -14.926810477657938: [],
 -26.901500884551737: [],
 7.335414115075125: [],
 37.269945214860485: [],
```

```
8.478143012041322: [],
17.482607274594955: [],
26.853584268838507: [],
11.76343558855115: [],
19.52645850327805: [],
11.63578332299442: [],
46.218006334531765: [],
32.66732554097734: [],
22.767291559664436: [],
3.9513452649045924: [],
2.6895773717807714: [],
-26.023436866434203: [],
39.861691980336055: [],
31.029575415168644: [],
-9.937908137060667: [],
-7.981643420780557: [],
-23.794798265137256: [],
9.235115167318568: [],
-21.20508963729302: [],
27.7627974661873: [],
7.91859270672829: [],
22.133868955902855: [],
28.876690635165218: [],
23.57986646122238: [],
-22.13341037493582: [],
4.6176810477657995: [],
10.04536894367402: [],
17.545216434754707: [],
24.47808090878879: ['gktrace'],
25.630457113507955: ['gktrace'],
19.661658962506436: ['gktrace'],
15.62777233039175: ['gktrace'],
16.65084562330324: ['gktrace'],
-12.872034231818539: ['gktrace'],
-0.6522855675398063: ['gkthighdegree'],
11.227208525937336: ['gkthighdegree'],
-3.474005592649661: ['gkthighdegree'],
22.818866632425966: ['gkthighdegree'],
23.81319245726025: ['gkthighdegree'],
1.18825971580209: ['gkthighdegree'],
-4.2501284026707715: ['gkthighdegree'],
-4.097212494747165: ['gkthighdegree'],
10.008237781735467: ['gkthighdegree'],
5.498691515640672: ['gkthighdegree'],
-5.142818788812619: ['gkthighdegree'],
-7.078839239856201: ['gkthighdegree'],
-5.877340430499132: ['gkthighdegree'],
-3.1785497501984423: ['gkthighdegree'],
4.908192090395481: ['gkthighdegree'],
4.036979969183363: ['gkthighdegree'],
8.326335727465672: ['gkthighdegree'],
-13.454130615147562: ['gkthighdegree'],
16.77393639787708: ['gkthighdegree'],
-6.255885122410547: ['gkthighdegree', 'gktrace'],
3.2773253112236205: ['gkschid'],
-0.3062403697996814: ['gkschid'],
-9.269856711324671: ['gkschid'],
```

```
-1.6517576898932802: ['gkschid'],
23.238855999783745: ['gkschid'],
-4.95022867904223: ['gkschid'],
-10.441582899210026: ['gkschid'],
-17.23067539804827: ['gkschid'],
-8.444867417948188: ['gkschid'],
-11.614250862132252: ['gkschid'],
-8.585318754810267: ['gkschid'],
15.314993673819622: ['gkschid'],
0.7813302516692318: ['gkschid'],
8.629243124620643: ['gkschid'],
-0.28586056128429504: ['gkschid'],
18.050190513730136: ['gkschid'],
-15.07237546619401: ['gkschid']}]
```

In [21]:
```
## Check the matched group with the most units in each trial -- also which cov
ariates did they use and which units in their MMG?
```

In [22]:
```
## Run DAME and FLAME and show why we chose DAME for this dataset. What happen
s if we run FLAME?
flame_models = []
random_seeds = [1111, 2222, 3333, 4444]
for i in range(4):
    matching_df, holdout_df = train_test_split(df_trunc, test_size=0.2, random
_state=random_seeds[i])
    model_flame = dame_flame.matching.FLAME(
        repeats=False, verbose=3, adaptive_weights='decisiontree',
        missing_holdout_replace=1, missing_data_replace=1)
    model_flame.fit(holdout_data=holdout_df)
    result_flame = model_flame.predict(matching_df)
    flame_models.append(model_flame)
```

```
Iteration number:  1
        Number of matched groups formed in total:  266
        Unmatched treated units:  443 out of a total of  1337 treated units
        Unmatched control units:  1381 out of a total of  3137 control units
        Predictive error of covariates chosen this iteration:  0
        Number of matches made in this iteration:  2650
        Number of matches made so far:  2650
        In this iteration, the covariates dropped are:  set()
Iteration number:  2
        Number of matched groups formed in total:  266
        Unmatched treated units:  443 out of a total of  1337 treated units
        Unmatched control units:  1381 out of a total of  3137 control units
        Predictive error of covariates chosen this iteration:  431.2160323161
8354
        Number of matches made in this iteration:  0
        Number of matches made so far:  2650
        In this iteration, the covariates dropped are:  gktgen
Iteration number:  3
        Number of matched groups formed in total:  337
        Unmatched treated units:  234 out of a total of  1337 treated units
        Unmatched control units:  904 out of a total of  3137 control units
        Predictive error of covariates chosen this iteration:  483.5169026371
4147
        Number of matches made in this iteration:  686
        Number of matches made so far:  3336
        In this iteration, the covariates dropped are:  gkthighdegree
Iteration number:  4
        Number of matched groups formed in total:  347
        Unmatched treated units:  209 out of a total of  1337 treated units
        Unmatched control units:  883 out of a total of  3137 control units
        Predictive error of covariates chosen this iteration:  547.1174748094
404
        Number of matches made in this iteration:  732
        Number of matches made so far:  3382
        In this iteration, the covariates dropped are:  race
Iteration number:  5
        Number of matched groups formed in total:  368
        Unmatched treated units:  103 out of a total of  1337 treated units
        Unmatched control units:  675 out of a total of  3137 control units
        Predictive error of covariates chosen this iteration:  613.0252062715
412
        Number of matches made in this iteration:  1046
        Number of matches made so far:  3696
        In this iteration, the covariates dropped are:  gktrace
Iteration number:  6
        Number of matched groups formed in total:  375
        Unmatched treated units:  90 out of a total of  1337 treated units
        Unmatched control units:  663 out of a total of  3137 control units
        Predictive error of covariates chosen this iteration:  784.8638930694
781
        Number of matches made in this iteration:  1071
        Number of matches made so far:  3721
        In this iteration, the covariates dropped are:  gender
Iteration number:  7
        Number of matched groups formed in total:  381
        Unmatched treated units:  81 out of a total of  1337 treated units
        Unmatched control units:  647 out of a total of  3137 control units
```

```
        Predictive error of covariates chosen this iteration:  949.4209481408
504
        Number of matches made in this iteration:  1096
        Number of matches made so far:  3746
        In this iteration, the covariates dropped are:  gkfreelunch
3746 units matched. No more covariate sets to consider dropping
Iteration number:  1
        Number of matched groups formed in total:  262
        Unmatched treated units:  439 out of a total of  1338 treated units
        Unmatched control units:  1394 out of a total of  3136 control units
        Predictive error of covariates chosen this iteration:  0
        Number of matches made in this iteration:  2641
        Number of matches made so far:  2641
        In this iteration, the covariates dropped are:  set()
Iteration number:  2
        Number of matched groups formed in total:  262
        Unmatched treated units:  439 out of a total of  1338 treated units
        Unmatched control units:  1394 out of a total of  3136 control units
        Predictive error of covariates chosen this iteration:  399.7143684669
0307
        Number of matches made in this iteration:  0
        Number of matches made so far:  2641
        In this iteration, the covariates dropped are:  gktgen
Iteration number:  3
        Number of matched groups formed in total:  283
        Unmatched treated units:  338 out of a total of  1338 treated units
        Unmatched control units:  1227 out of a total of  3136 control units
        Predictive error of covariates chosen this iteration:  429.6920147194
0426
        Number of matches made in this iteration:  268
        Number of matches made so far:  2909
        In this iteration, the covariates dropped are:  gktrace
Iteration number:  4
        Number of matched groups formed in total:  290
        Unmatched treated units:  330 out of a total of  1338 treated units
        Unmatched control units:  1202 out of a total of  3136 control units
        Predictive error of covariates chosen this iteration:  465.2032575934
312
        Number of matches made in this iteration:  301
        Number of matches made so far:  2942
        In this iteration, the covariates dropped are:  race
Iteration number:  5
        Number of matched groups formed in total:  359
        Unmatched treated units:  93 out of a total of  1338 treated units
        Unmatched control units:  688 out of a total of  3136 control units
        Predictive error of covariates chosen this iteration:  557.8764471699
633
        Number of matches made in this iteration:  1052
        Number of matches made so far:  3693
        In this iteration, the covariates dropped are:  gkthighdegree
Iteration number:  6
        Number of matched groups formed in total:  362
        Unmatched treated units:  89 out of a total of  1338 treated units
        Unmatched control units:  683 out of a total of  3136 control units
        Predictive error of covariates chosen this iteration:  746.0154567540
561
        Number of matches made in this iteration:  1061
```

```
        Number of matches made so far:  3702
        In this iteration, the covariates dropped are:  gender
Iteration number:  7
        Number of matched groups formed in total:  366
        Unmatched treated units:  79 out of a total of  1338 treated units
        Unmatched control units:  653 out of a total of  3136 control units
        Predictive error of covariates chosen this iteration:  910.1537265767
727
        Number of matches made in this iteration:  1101
        Number of matches made so far:  3742
        In this iteration, the covariates dropped are:  gkfreelunch
3742 units matched. No more covariate sets to consider dropping
Iteration number:  1
        Number of matched groups formed in total:  263
        Unmatched treated units:  431 out of a total of  1355 treated units
        Unmatched control units:  1419 out of a total of  3119 control units
        Predictive error of covariates chosen this iteration:  0
        Number of matches made in this iteration:  2624
        Number of matches made so far:  2624
        In this iteration, the covariates dropped are:  set()
Iteration number:  2
        Number of matched groups formed in total:  263
        Unmatched treated units:  431 out of a total of  1355 treated units
        Unmatched control units:  1419 out of a total of  3119 control units
        Predictive error of covariates chosen this iteration:  388.0985821238
4574
        Number of matches made in this iteration:  0
        Number of matches made so far:  2624
        In this iteration, the covariates dropped are:  gktgen
Iteration number:  3
        Number of matched groups formed in total:  282
        Unmatched treated units:  344 out of a total of  1355 treated units
        Unmatched control units:  1274 out of a total of  3119 control units
        Predictive error of covariates chosen this iteration:  422.1180885299
973
        Number of matches made in this iteration:  232
        Number of matches made so far:  2856
        In this iteration, the covariates dropped are:  gktrace
Iteration number:  4
        Number of matched groups formed in total:  289
        Unmatched treated units:  333 out of a total of  1355 treated units
        Unmatched control units:  1249 out of a total of  3119 control units
        Predictive error of covariates chosen this iteration:  473.5321698961
1747
        Number of matches made in this iteration:  268
        Number of matches made so far:  2892
        In this iteration, the covariates dropped are:  race
Iteration number:  5
        Number of matched groups formed in total:  367
        Unmatched treated units:  80 out of a total of  1355 treated units
        Unmatched control units:  705 out of a total of  3119 control units
        Predictive error of covariates chosen this iteration:  533.4808904871
547
        Number of matches made in this iteration:  1065
        Number of matches made so far:  3689
        In this iteration, the covariates dropped are:  gkthighdegree
Iteration number:  6
```

```
        Number of matched groups formed in total:  371
        Unmatched treated units:  74 out of a total of  1355 treated units
        Unmatched control units:  683 out of a total of  3119 control units
        Predictive error of covariates chosen this iteration:  708.4432689675
247
        Number of matches made in this iteration:  1093
        Number of matches made so far:  3717
        In this iteration, the covariates dropped are:  gkfreelunch
Iteration number:  7
        Number of matched groups formed in total:  373
        Unmatched treated units:  71 out of a total of  1355 treated units
        Unmatched control units:  680 out of a total of  3119 control units
        Predictive error of covariates chosen this iteration:  885.3999471421
853
        Number of matches made in this iteration:  1099
        Number of matches made so far:  3723
        In this iteration, the covariates dropped are:  gender
3723 units matched. No more covariate sets to consider dropping
Iteration number:  1
        Number of matched groups formed in total:  264
        Unmatched treated units:  435 out of a total of  1383 treated units
        Unmatched control units:  1403 out of a total of  3091 control units
        Predictive error of covariates chosen this iteration:  0
        Number of matches made in this iteration:  2636
        Number of matches made so far:  2636
        In this iteration, the covariates dropped are:  set()
Iteration number:  2
        Number of matched groups formed in total:  264
        Unmatched treated units:  435 out of a total of  1383 treated units
        Unmatched control units:  1403 out of a total of  3091 control units
        Predictive error of covariates chosen this iteration:  404.4129879605
2354
        Number of matches made in this iteration:  0
        Number of matches made so far:  2636
        In this iteration, the covariates dropped are:  gktgen
Iteration number:  3
        Number of matched groups formed in total:  286
        Unmatched treated units:  342 out of a total of  1383 treated units
        Unmatched control units:  1255 out of a total of  3091 control units
        Predictive error of covariates chosen this iteration:  432.4330990941
5893
        Number of matches made in this iteration:  241
        Number of matches made so far:  2877
        In this iteration, the covariates dropped are:  gktrace
Iteration number:  4
        Number of matched groups formed in total:  292
        Unmatched treated units:  335 out of a total of  1383 treated units
        Unmatched control units:  1236 out of a total of  3091 control units
        Predictive error of covariates chosen this iteration:  492.9382058799
0126
        Number of matches made in this iteration:  267
        Number of matches made so far:  2903
        In this iteration, the covariates dropped are:  race
Iteration number:  5
        Number of matched groups formed in total:  364
        Unmatched treated units:  89 out of a total of  1383 treated units
        Unmatched control units:  713 out of a total of  3091 control units
```

```
        Predictive error of covariates chosen this iteration:  573.7005294998
937
        Number of matches made in this iteration:  1036
        Number of matches made so far:  3672
        In this iteration, the covariates dropped are:  gkthighdegree
Iteration number:  6
        Number of matched groups formed in total:  370
        Unmatched treated units:  81 out of a total of  1383 treated units
        Unmatched control units:  704 out of a total of  3091 control units
        Predictive error of covariates chosen this iteration:  706.8890879460
384
        Number of matches made in this iteration:  1053
        Number of matches made so far:  3689
        In this iteration, the covariates dropped are:  gender
Iteration number:  7
        Number of matched groups formed in total:  374
        Unmatched treated units:  70 out of a total of  1383 treated units
        Unmatched control units:  689 out of a total of  3091 control units
        Predictive error of covariates chosen this iteration:  873.1535531789
737
        Number of matches made in this iteration:  1079
        Number of matches made so far:  3715
        In this iteration, the covariates dropped are:  gkfreelunch
3715 units matched. No more covariate sets to consider dropping
```

In [23]:
```
# whats the var of the ates?
for model in flame_models:
    print(dame_flame.utils.post_processing.var_ATE(matching_object=model))
```

```
treated_col treated
(5.613833283201907, 0.8748410806557086)
treated_col treated
(4.712491220580468, 0.8402015032500467)
treated_col treated
(5.287778786150933, 0.8443983228179541)
treated_col treated
(5.836330481668296, 0.8402137371729708)
```

In [ ]:
```

```