CME_Tick_Changes_JPY

October 9, 2019

1 The Robert and Rosenbaum Uncertainty Zones model

2 An application to EURUSD FX Futures at CME

2.1 Implementation by

2.2 Marcos Costa Santos Carreira (École Polytechnique - CMAP)

2.3 and

2.4 Florian Huchedé (CME)

2.5 Aug-2019

2.6 Import packages

[1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import statsmodels.api as sm
import glob
```

[2]:
```
pd.set_option('display.max_columns', 50)
```

[3]:
```
pd.set_option('display.max_rows', 200)
```

[4]:
```
import cme_processing as cme
```

2.7 File paths and initial values

[5]:
```
PATHPROJ = '/Users/marcoscscarreira/Documents/X/CME project/CME_data/'
URL_ROOT = 'https://raw.githubusercontent.com/MarcosCarreira/UZStats/master/
 ↪CME_data/'
```

[6]:
```
CURR = 'JPY'
```

[7]:
```
PATH_PRIOR = PATHPROJ+CURR+'/prior/'
PATH_AFTER = PATHPROJ+CURR+'/after/'
URL_1 = CURR+'/prior/'
URL_2 = CURR+'/after/'
#PATH_PRIOR = URL_ROOT+URL_1
#PATH_AFTER = URL_ROOT+URL_2
```

[8]:
```
TRADING_HOURS = 8
```

[9]:
```
TICK_PRIOR = 1.0
TICK_AFTER = 0.5
```

[10]:
```
PRIOR_CDATES_LIST = [['6JH5', '20150105'], ['6JH5', '20150106'], ['6JH5',␣
↪'20150107'],\
   ['6JH5', '20150108'], ['6JH5', '20150109'], ['6JH5', '20150112'], ['6JH5',␣
↪'20150113'],\
   ['6JH5', '20150114'], ['6JH5', '20150115'], ['6JH5', '20150116'], ['6JH5',␣
↪'20150119'],\
   ['6JH5', '20150120'], ['6JH5', '20150121'], ['6JH5', '20150122'], ['6JH5',␣
↪'20150123'],\
   ['6JH5', '20150126'], ['6JH5', '20150127'], ['6JH5', '20150128'], ['6JH5',␣
↪'20150129'],\
   ['6JH5', '20150130'], ['6JH5', '20150202'], ['6JH5', '20150203'], ['6JH5',␣
↪'20150204'],\
   ['6JH5', '20150205'], ['6JH5', '20150206'], ['6JH5', '20150209'], ['6JH5',␣
↪'20150210'],\
   ['6JH5', '20150211'], ['6JH5', '20150212'], ['6JH5', '20150213'], ['6JH5',␣
↪'20150223'],\
   ['6JH5', '20150224'], ['6JH5', '20150225'], ['6JH5', '20150226'], ['6JH5',␣
↪'20150227'],\
   ['6JH5', '20150302'], ['6JH5', '20150303'], ['6JH5', '20150304'], ['6JH5',␣
↪'20150305'],\
   ['6JH5', '20150306'], ['6JH5', '20150309'], ['6JH5', '20150310'], ['6JH5',␣
↪'20150311'],\
   ['6JH5', '20150312'], ['6JH5', '20150313'], ['6JM5', '20150316'], ['6JM5',␣
↪'20150317'],\
   ['6JM5', '20150318'], ['6JM5', '20150319'], ['6JM5', '20150320'], ['6JM5',␣
↪'20150323'],\
   ['6JM5', '20150324'], ['6JM5', '20150325'], ['6JM5', '20150326'], ['6JM5',␣
↪'20150327'],\
   ['6JM5', '20150330'], ['6JM5', '20150331'], ['6JM5', '20150401'], ['6JM5',␣
↪'20150402'],\
   ['6JM5', '20150403'], ['6JM5', '20150406'], ['6JM5', '20150407'], ['6JM5',␣
↪'20150408'],\
```

```
   ['6JM5', '20150409'], ['6JM5', '20150410'], ['6JM5', '20150413'], ['6JM5',␣
↪'20150414'],\
   ['6JM5', '20150415'], ['6JM5', '20150416'], ['6JM5', '20150417'], ['6JM5',␣
↪'20150420'],\
   ['6JM5', '20150421'], ['6JM5', '20150422'], ['6JM5', '20150423'], ['6JM5',␣
↪'20150424'],\
   ['6JM5', '20150427'], ['6JM5', '20150428'], ['6JM5', '20150429'], ['6JM5',␣
↪'20150430'],\
   ['6JM5', '20150501'], ['6JM5', '20150504'], ['6JM5', '20150505'], ['6JM5',␣
↪'20150506'],\
   ['6JM5', '20150507'], ['6JM5', '20150508'], ['6JM5', '20150511'], ['6JM5',␣
↪'20150512'],\
   ['6JM5', '20150513'], ['6JM5', '20150514'], ['6JM5', '20150515'], ['6JM5',␣
↪'20150518'],\
   ['6JM5', '20150519'], ['6JM5', '20150520'], ['6JM5', '20150521'], ['6JM5',␣
↪'20150522'],\
   ['6JM5', '20150525'], ['6JM5', '20150526'], ['6JM5', '20150527'], ['6JM5',␣
↪'20150528'],\
   ['6JM5', '20150529'], ['6JM5', '20150601'], ['6JM5', '20150602'], ['6JM5',␣
↪'20150603'],\
   ['6JM5', '20150604'], ['6JM5', '20150605'], ['6JM5', '20150608'], ['6JM5',␣
↪'20150609'],\
   ['6JM5', '20150610'], ['6JM5', '20150611'], ['6JU5', '061215'], ['6JU5',␣
↪'061515'],\
   ['6JU5', '061615'], ['6JU5', '061715'], ['6JU5', '061815'], ['6JU5',␣
↪'061915']]
```

[11]:
```
AFTER_CDATES_LIST = [['6JU5', '062215'], ['6JU5', '062315'], ['6JU5',␣
↪'062415'],\
   ['6JU5', '062515'], ['6JU5', '062615'], ['6JU5', '062915'], ['6JU5',␣
↪'063015'],\
   ['6JU5', '070115'], ['6JU5', '070215'], ['6JU5', '070315'], ['6JU5',␣
↪'070615'],\
   ['6JU5', '070715'], ['6JU5', '070815'], ['6JU5', '070915'], ['6JU5',␣
↪'071015'],\
   ['6JU5', '20150713'], ['6JU5', '20150714'], ['6JU5', '20150715'], ['6JU5',␣
↪'20150716'],\
   ['6JU5', '20150717'], ['6JU5', '20150720'], ['6JU5', '20150721'], ['6JU5',␣
↪'20150722'],\
   ['6JU5', '20150723'], ['6JU5', '20150724'], ['6JU5', '20150727'], ['6JU5',␣
↪'20150728'],\
   ['6JU5', '20150729'], ['6JU5', '20150730'], ['6JU5', '20150731'], ['6JU5',␣
↪'20150803'],\
   ['6JU5', '20150804'], ['6JU5', '20150805'], ['6JU5', '20150806'], ['6JU5',␣
↪'20150807'],\
```

```
   ['6JU5', '20150810'], ['6JU5', '20150811'], ['6JU5', '20150812'], ['6JU5',␣
↪'20150813'],\
   ['6JU5', '20150814'], ['6JU5', '20150817'], ['6JU5', '20150818'], ['6JU5',␣
↪'20150819'],\
   ['6JU5', '20150820'], ['6JU5', '20150821'], ['6JU5', '20150824'], ['6JU5',␣
↪'20150825'],\
   ['6JU5', '20150826'], ['6JU5', '20150827'], ['6JU5', '20150828'], ['6JU5',␣
↪'20150831'],\
   ['6JU5', '20150901'], ['6JU5', '20150902'], ['6JU5', '20150903'], ['6JU5',␣
↪'20150904'],\
   ['6JU5', '20150907'], ['6JU5', '20150908'], ['6JU5', '20150909'], ['6JU5',␣
↪'20150910'],\
   ['6JU5', '20150911'], ['6JZ5', '20150914'], ['6JZ5', '20150915'], ['6JZ5',␣
↪'20150916'],\
   ['6JZ5', '20150917'], ['6JZ5', '20150918'], ['6JZ5', '20150921'], ['6JZ5',␣
↪'20150922'],\
   ['6JZ5', '20150923'], ['6JZ5', '20150924'], ['6JZ5', '20150925'], ['6JZ5',␣
↪'20150928'],\
   ['6JZ5', '20150929'], ['6JZ5', '20150930'], ['6JZ5', '20151001'], ['6JZ5',␣
↪'20151002'],\
   ['6JZ5', '20151005'], ['6JZ5', '20151006'], ['6JZ5', '20151007'], ['6JZ5',␣
↪'20151008'],\
   ['6JZ5', '20151009'], ['6JZ5', '20151012'], ['6JZ5', '20151013'], ['6JZ5',␣
↪'20151014'],\
   ['6JZ5', '20151015'], ['6JZ5', '20151016'], ['6JZ5', '20151019'], ['6JZ5',␣
↪'20151020'],\
   ['6JZ5', '20151021'], ['6JZ5', '20151022'], ['6JZ5', '20151023'], ['6JZ5',␣
↪'20151026'],\
   ['6JZ5', '20151027'], ['6JZ5', '20151028'], ['6JZ5', '20151029'], ['6JZ5',␣
↪'20151030'],\
   ['6JZ5', '20151102'], ['6JZ5', '20151103'], ['6JZ5', '20151104'], ['6JZ5',␣
↪'20151105'],\
   ['6JZ5', '20151106'], ['6JZ5', '20151109'], ['6JZ5', '20151110'], ['6JZ5',␣
↪'20151111'],\
   ['6JZ5', '20151112'], ['6JZ5', '20151113'], ['6JZ5', '20151116'], ['6JZ5',␣
↪'20151117'],\
   ['6JZ5', '20151118'], ['6JZ5', '20151119'], ['6JZ5', '20151120'], ['6JZ5',␣
↪'20151123'],\
   ['6JZ5', '20151124'], ['6JZ5', '20151125'], ['6JZ5', '20151126'], ['6JZ5',␣
↪'20151127'],\
   ['6JZ5', '20151130'], ['6JZ5', '20151201'], ['6JZ5', '20151202'], ['6JZ5',␣
↪'20151203'],\
   ['6JZ5', '20151204'], ['6JZ5', '20151207'], ['6JZ5', '20151208'], ['6JZ5',␣
↪'20151209'],\
   ['6JZ5', '20151210'], ['6JZ5', '20151211']]
```

2.7.1 Processing files

Prior

[12]:
```
#PRIOR_CDATES_LIST = cme.list_files(PATH_PRIOR)
```

[13]:
```
#PRIOR_CDATES_LIST
```

[14]:
```
PRIOR_CDATES, FILES_PRIOR_CAticks, FILES_PRIOR_COSTtrades,\
    FILES_PRIOR_OBstats, FILES_PRIOR_OTtrans,\
    FILES_PRIOR_RDFtrans, FILES_PRIOR_UZstats = \
    cme.process_files(PATH_PRIOR, PRIOR_CDATES_LIST, 'prior', TICK_PRIOR)
```

[15]:
```
PRIOR_OB_UZ_STATS = cme.ob_uz_stats(PRIOR_CDATES, FILES_PRIOR_OBstats,\
    FILES_PRIOR_UZstats, FILES_PRIOR_CAticks, TRADING_HOURS)
```

[16]:
```
PRIOR_IMBAL_STATS = cme.imbal_stats(PRIOR_CDATES, FILES_PRIOR_OTtrans)
```

[17]:
```
PRIOR_IMBAL_STATS_TS = cme.time_series_imbal(PRIOR_IMBAL_STATS, pd.
↪to_datetime(PRIOR_CDATES['Date']), 'prior')
```

[18]:
```
PRIOR_IMBAL_STATS_TS['eta1'] = PRIOR_OB_UZ_STATS['eta1'].values
```

[19]:
```
PRIOR_TRADE_STATS_TS = cme.time_series_imbal_trd(PRIOR_IMBAL_STATS, pd.
↪to_datetime(PRIOR_CDATES['Date']), 'prior')
```

[20]:
```
PRIOR_DEPL_STATS = cme.depl_stats(PRIOR_CDATES, FILES_PRIOR_RDFtrans)
```

[21]:
```
PRIOR_DEPL_STATS_TS = cme.time_series_depl(PRIOR_DEPL_STATS, pd.
↪to_datetime(PRIOR_CDATES['Date']), 'prior')
```

[22]:
```
PRIOR_DEPL_STATS_TS['eta1'] = PRIOR_OB_UZ_STATS['eta1'].values
```

[23]:
```
PRIOR_ABSDEPL_STATS_TS = cme.time_series_absdepl(PRIOR_DEPL_STATS, pd.
↪to_datetime(PRIOR_CDATES['Date']), 'prior')
```

[24]:
```
PRIOR_ABSDEPL_STATS_TS['eta1'] = PRIOR_OB_UZ_STATS['eta1'].values
PRIOR_ABSDEPL_STATS_TS['M'] = PRIOR_OB_UZ_STATS['M'].values
```

[25]:
```
PRIOR_COST_STATS = cme.cost_stats(PRIOR_CDATES, FILES_PRIOR_COSTtrades)
```

[26]:
```
PRIOR_COST_STATS['Status'] = 'prior'
```

After

[27]:
```
#AFTER_CDATES_LIST = cme.list_files(PATH_AFTER)
```

[28]:
```
#AFTER_CDATES_LIST
```

[29]:
```
AFTER_CDATES, FILES_AFTER_CAticks, FILES_AFTER_COSTtrades,\
    FILES_AFTER_OBstats, FILES_AFTER_OTtrans,\
    FILES_AFTER_RDFtrans, FILES_AFTER_UZstats = \
    cme.process_files(PATH_AFTER, AFTER_CDATES_LIST, 'after', TICK_AFTER)
```

[30]:
```
AFTER_OB_UZ_STATS = cme.ob_uz_stats(AFTER_CDATES, FILES_AFTER_OBstats,\
    FILES_AFTER_UZstats, FILES_AFTER_CAticks, TRADING_HOURS)
```

[31]:
```
AFTER_IMBAL_STATS = cme.imbal_stats(AFTER_CDATES, FILES_AFTER_OTtrans)
```

[32]:
```
AFTER_IMBAL_STATS_TS = cme.time_series_imbal(AFTER_IMBAL_STATS, pd.
↪to_datetime(AFTER_CDATES['Date']), 'after')
```

[33]:
```
AFTER_IMBAL_STATS_TS['eta1'] = AFTER_OB_UZ_STATS['eta1'].values
```

[34]:
```
AFTER_TRADE_STATS_TS = cme.time_series_imbal_trd(AFTER_IMBAL_STATS, pd.
↪to_datetime(AFTER_CDATES['Date']), 'after')
```

[35]:
```
AFTER_DEPL_STATS = cme.depl_stats(AFTER_CDATES, FILES_AFTER_RDFtrans)
```

[36]:
```
AFTER_DEPL_STATS_TS = cme.time_series_depl(AFTER_DEPL_STATS, pd.
↪to_datetime(AFTER_CDATES['Date']), 'after')
```

[37]:
```
AFTER_DEPL_STATS_TS['eta1'] = AFTER_OB_UZ_STATS['eta1'].values
```

[38]:
```
AFTER_ABSDEPL_STATS_TS = cme.time_series_absdepl(AFTER_DEPL_STATS, pd.
↪to_datetime(AFTER_CDATES['Date']), 'after')
```

[39]:
```
AFTER_ABSDEPL_STATS_TS['eta1'] = AFTER_OB_UZ_STATS['eta1'].values
AFTER_ABSDEPL_STATS_TS['M'] = AFTER_OB_UZ_STATS['M'].values
```

[40]:
```
AFTER_COST_STATS = cme.cost_stats(AFTER_CDATES, FILES_AFTER_COSTtrades)
```

[41]:
```
AFTER_COST_STATS['Status'] = 'after'
```

Join prior and after

[42]:
```
OB_UZ_STATS = pd.concat([PRIOR_OB_UZ_STATS, AFTER_OB_UZ_STATS], sort=False)
```

[43]:
```
IMBAL_STATS_TS = pd.concat([PRIOR_IMBAL_STATS_TS, AFTER_IMBAL_STATS_TS],␣
↪sort=False)
```

[44]:
```
TRADE_STATS_TS = pd.concat([PRIOR_TRADE_STATS_TS, AFTER_TRADE_STATS_TS],␣
↪sort=False)
```

[45]:
```
DEPL_STATS_TS = pd.concat([PRIOR_DEPL_STATS_TS, AFTER_DEPL_STATS_TS],␣
↪sort=False)
```

[46]:
```
ABSDEPL_STATS_TS = pd.concat([PRIOR_ABSDEPL_STATS_TS, AFTER_ABSDEPL_STATS_TS],␣
↪sort=False)
```

2.7.2 Tables

[47]:
```
TABLE_MATHIEU = cme.table_mathieu(OB_UZ_STATS)
TABLE_MATHIEU_ERR = cme.table_mathieu_err(OB_UZ_STATS)
```

[48]:
```
TABLE_MATHIEU
```

[48]:

Status	Tick	chgavg	ndfpr_pred	ndfpr	M	Volume
prior	1.0	1.02531	2539.57754	1993.7913	13109.66957	75510.02609
after	0.5	0.52336	6152.96473	3809.2800	17209.84000	68863.32000

Status	eta1	S1	lambda1	twspr1	duration	dt_avg	rvxe
prior	0.23994	0.98738	0.98544	1.04473	16.91766	17.73267	0.00400
after	0.33870	0.96178	0.97067	1.16956	9.82449	10.58436	0.00348

Status	spot_avg
prior	8330.35714
after	8205.53733

[49]:
```
TABLE_MATHIEU_ERR
```

[49]:

Status	Tick	chgavg	ndfpr_pred	ndfpr	M	Volume
prior	0.0	0.04983	2079.37315	1024.36377	6368.99859	32942.32568
after	0.0	0.03153	10626.03777	3207.30554	9190.30675	34884.79666

Status	eta1	S1	lambda1	twspr1	duration	dt_avg	rvxe
prior	0.05823	0.01332	0.01780	0.04863	11.28649	8.49775	0.00175
after	0.06312	0.03516	0.03298	0.12921	9.44989	6.71907	0.00206

Status	spot_avg
prior	135.75089
after	119.74845

[50]:
```
cme.avg_perc_mat(PRIOR_IMBAL_STATS, pd.to_datetime(PRIOR_CDATES['Date']))
```

[50]:

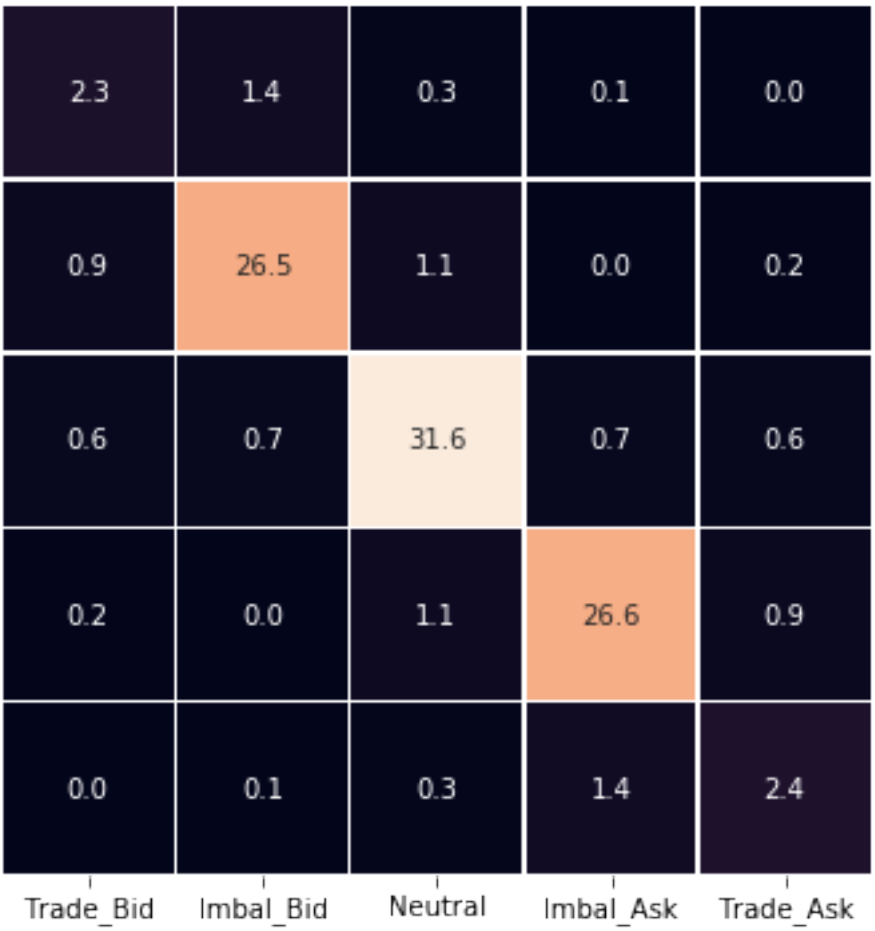
	Trade_Bid	Imbal_Bid	Neutral	Imbal_Ask	Trade_Ask	Total Cols
Trade_Bid	2.33	1.39	0.27	0.08	0.00	4.07
Imbal_Bid	0.88	26.48	1.09	0.04	0.23	28.72
Neutral	0.62	0.74	31.63	0.73	0.60	34.32
Imbal_Ask	0.23	0.04	1.08	26.57	0.88	28.81
Trade_Ask	0.00	0.07	0.26	1.38	2.37	4.08
Total Rows	4.07	28.72	34.32	28.81	4.08	100.00

[51]:
```
cme.avg_perc_mat(AFTER_IMBAL_STATS, pd.to_datetime(AFTER_CDATES['Date']))
```

[51]:

	Trade_Bid	Imbal_Bid	Neutral	Imbal_Ask	Trade_Ask	Total Cols
Trade_Bid	0.91	1.06	0.28	0.09	0.00	2.34
Imbal_Bid	0.61	24.46	1.64	0.16	0.24	27.11
Neutral	0.58	1.35	37.09	1.35	0.57	40.94
Imbal_Ask	0.25	0.16	1.64	24.61	0.60	27.25
Trade_Ask	0.00	0.09	0.28	1.05	0.93	2.35
Total Rows	2.34	27.11	40.94	27.25	2.35	100.00

[52]:
```
AVG_IMBAL_PRIOR = cme.avg_perc_mat(PRIOR_IMBAL_STATS, pd.
 ↪to_datetime(PRIOR_CDATES['Date']))
plt.figure(figsize=(9, 6))
sns.heatmap(AVG_IMBAL_PRIOR.iloc[:-1].drop(columns=['Total Cols']),\
    annot=True, fmt=".1f",\
    linewidths=.5, square=True,\
    xticklabels=True,\
    yticklabels=False,\
    cbar=False);
```



[53]:
```
cme.avg_perc_mat_2(PRIOR_DEPL_STATS, pd.to_datetime(PRIOR_CDATES['Date']))
```

[53]:

	same				oppo				Total Cols
	D C	D T	D T+F	F	D C	D T	D T+F	F	
D C	0.02	0.03	0.11	17.36	0.02	0.01	2.32	0.59	20.48
D T	0.01	0.11	0.52	18.07	0.02	0.02	3.81	2.72	25.30
D T+F	0.01	0.02	0.24	5.77	0.01	0.01	1.54	0.95	8.55
F	12.54	8.70	0.00	0.10	7.84	16.40	0.00	0.10	45.67
Total Rows	12.59	8.86	0.88	41.30	7.89	16.44	7.68	4.37	100.00

[54]:
```
cme.avg_perc_mat_2(AFTER_DEPL_STATS, pd.to_datetime(AFTER_CDATES['Date']))
```

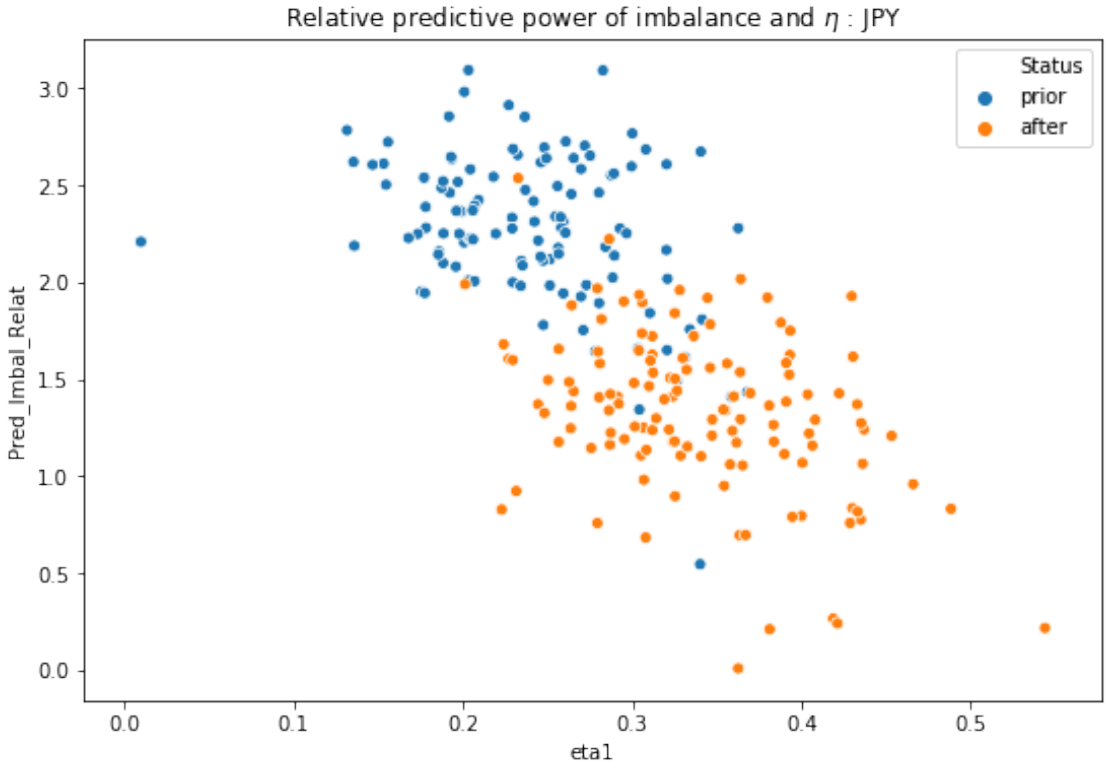
[54]:

	same				oppo				Total Cols
	D C	D T	D T+F	F	D C	D T	D T+F	F	
D C	0.04	0.05	0.15	23.48	0.03	0.02	1.24	1.49	26.51
D T	0.02	0.14	0.49	13.94	0.05	0.04	2.28	4.14	21.08
D T+F	0.02	0.03	0.19	3.12	0.02	0.02	0.61	0.97	4.96
F	19.04	7.43	0.00	0.16	7.29	13.37	0.00	0.15	47.44

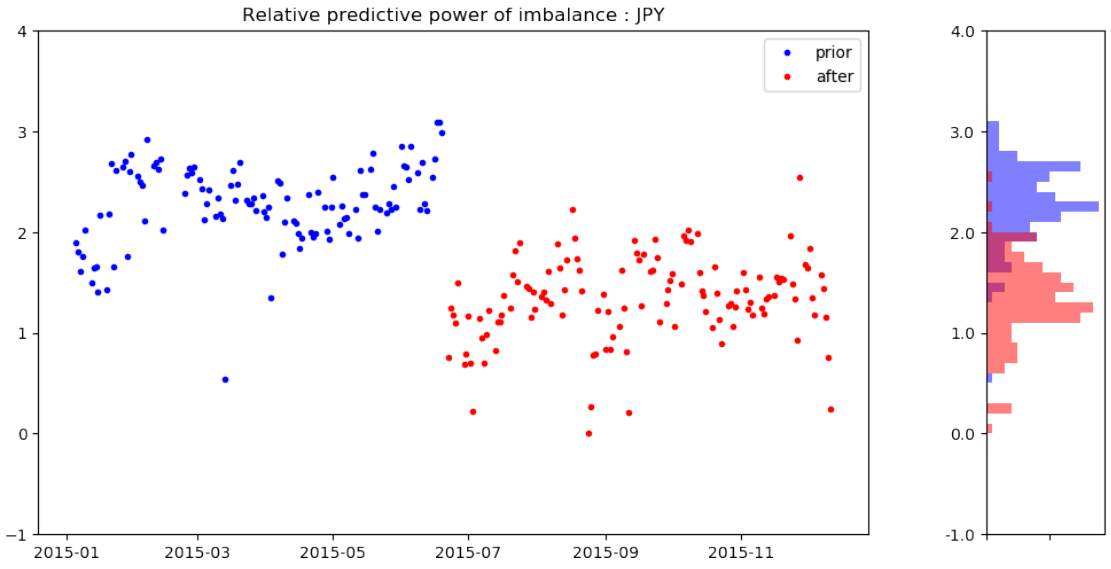
```
Total Rows  19.13  7.65  0.83  40.69  7.38  13.44  4.13  6.75     100.00
```

2.8 Charts and Regressions

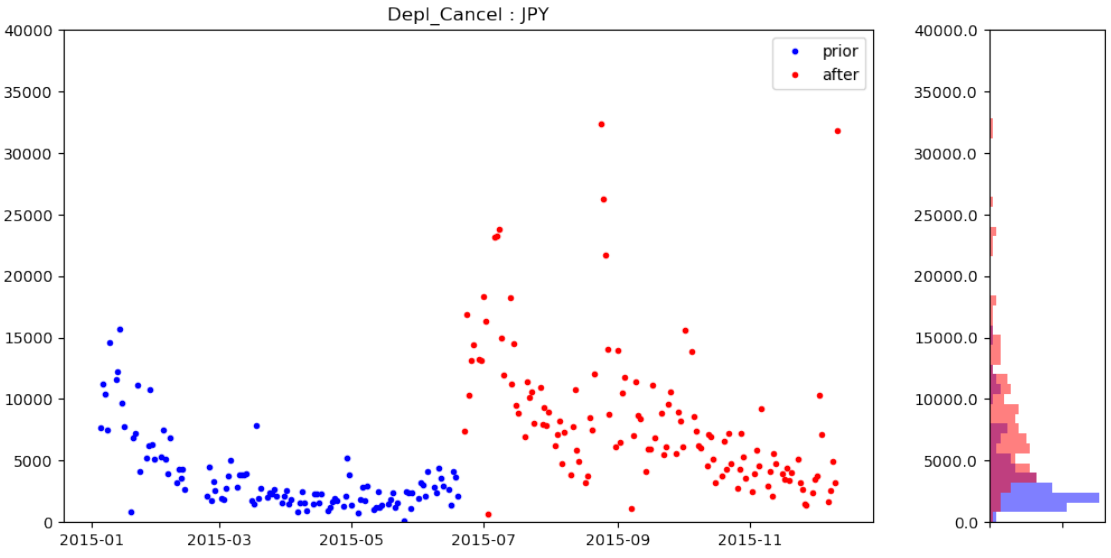
[55]:
```
plt.figure(figsize=(9, 6))
sns.scatterplot(x='eta1', y='Pred_Imbal_Relat', hue='Status',\
           data=IMBAL_STATS_TS);
plt.title('Relative predictive power of imbalance and $\eta$ : '+CURR);
```



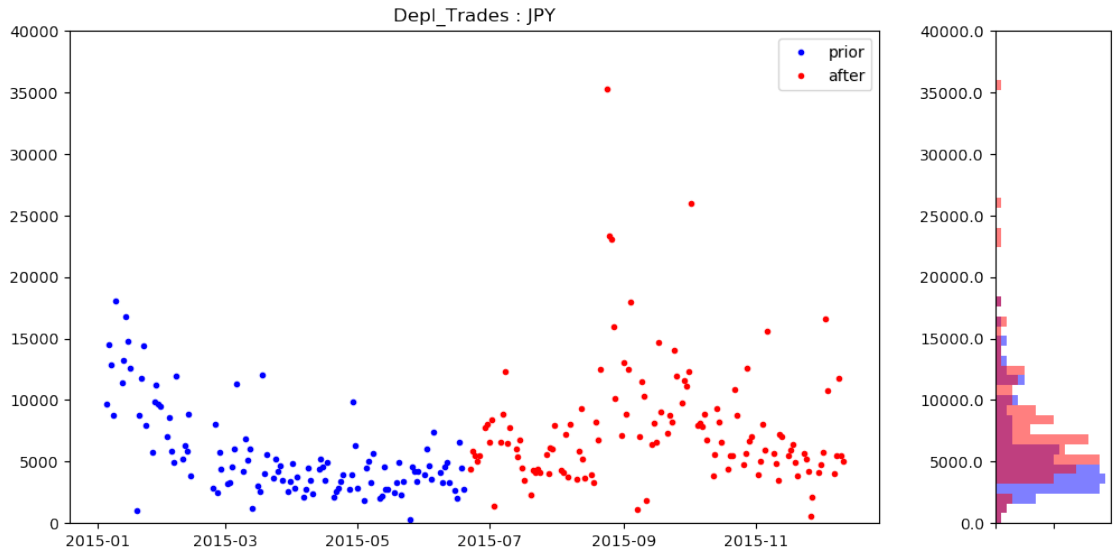
[164]:
```
cme.time_series_hist_plot(IMBAL_STATS_TS, 'Pred_Imbal_Relat',\
    'Relative predictive power of imbalance : '+CURR, -1.0, 4.0, 50)
```



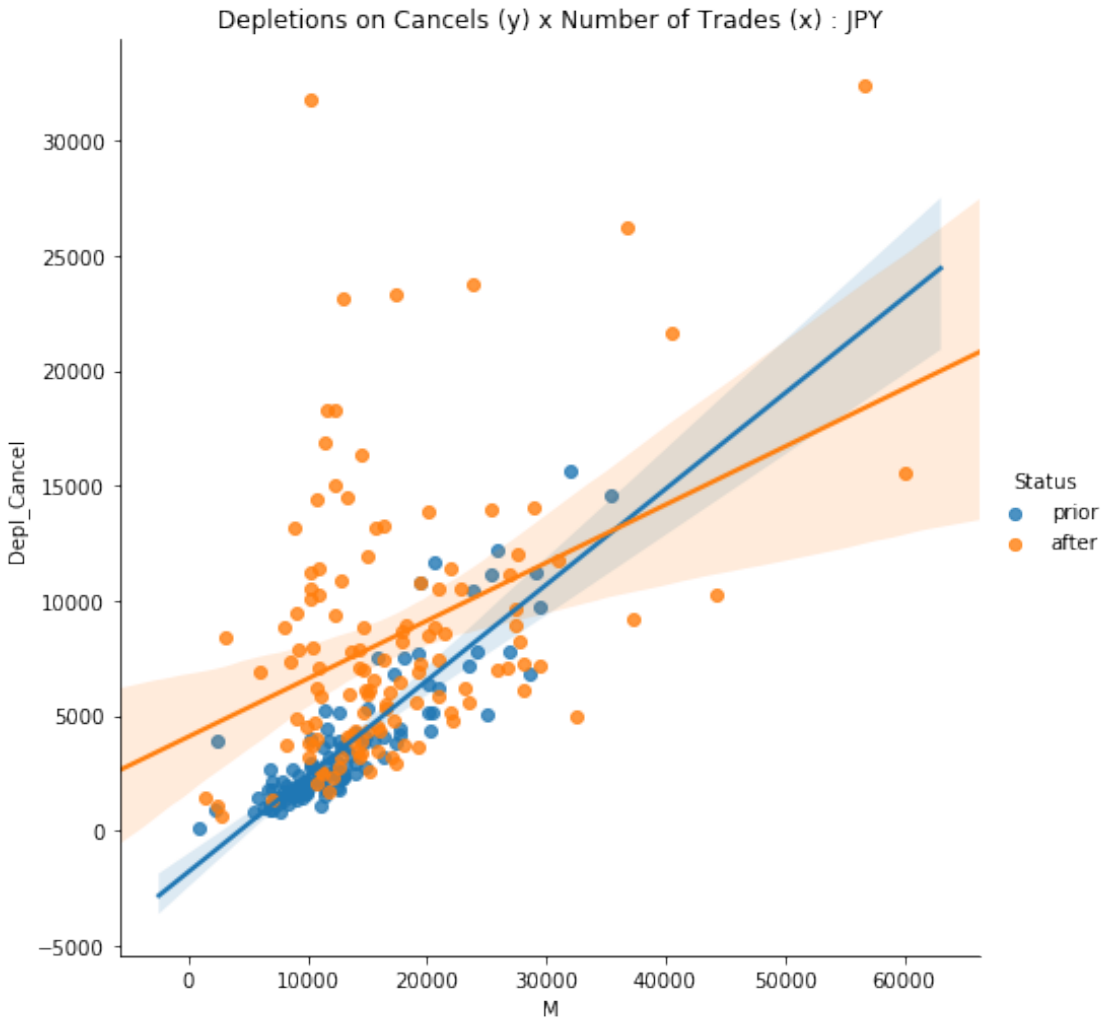
[165]:
```
cme.time_series_hist_plot(ABSDEPL_STATS_TS, 'Depl_Cancel',\
    'Depl_Cancel : '+CURR, 0.0, 40000.0, 50)
```



[166]:
```
cme.time_series_hist_plot(ABSDEPL_STATS_TS, 'Depl_Trades',\
    'Depl_Trades : '+CURR, 0.0, 40000.0, 50)
```



[59]:
```
cme.regr_plot(ABSDEPL_STATS_TS, 'M', 'Depl_Cancel',\
    'Depletions on Cancels (y) x Number of Trades (x) : '+CURR)
```



[60]:
```
cme.regr_plot(ABSDEPL_STATS_TS, 'M', 'Depl_Trades',\
    'Depletions on Trades (y) x Number of Trades (x) : '+CURR)
```

[61]:
```
cme.lin_reg(ABSDEPL_STATS_TS, ['M'], 'Depl_Cancel')
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:            Depl_Cancel   R-squared:                       0.309
Model:                            OLS   Adj. R-squared:                  0.306
Method:                 Least Squares   F-statistic:                     106.2
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           7.88e-21
Time:                        16:10:27   Log-Likelihood:                -2352.2
No. Observations:                 240   AIC:                             4708.
Df Residuals:                     238   BIC:                             4715.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

```
const        714.4925    598.043      1.195      0.233    -463.640    1892.625
M              0.3561      0.035     10.305      0.000       0.288       0.424
==============================================================================
Omnibus:                      146.578   Durbin-Watson:                   0.608
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              944.074
Skew:                           2.452   Prob(JB):                    9.93e-206
Kurtosis:                      11.388   Cond. No.                     3.66e+04
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 3.66e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

```
/Users/marcoscscarreira/anaconda3/envs/CondaEnv36/lib/python3.6/site-
packages/numpy/core/fromnumeric.py:2495: FutureWarning: Method .ptp is
deprecated and will be removed in a future version. Use numpy.ptp instead.
  return ptp(axis=axis, out=out, **kwargs)
```

[62]:
```
cme.lin_reg(ABSDEPL_STATS_TS, ['M'], 'Depl_Trades')
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:            Depl_Trades   R-squared:                       0.915
Model:                            OLS   Adj. R-squared:                  0.915
Method:                 Least Squares   F-statistic:                     2558.
Date:                Wed, 09 Oct 2019   Prob (F-statistic):          2.56e-129
Time:                        16:10:27   Log-Likelihood:                -2059.3
No. Observations:                 240   AIC:                             4123.
Df Residuals:                     238   BIC:                             4130.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const      -1183.8240    176.517     -6.707      0.000   -1531.560    -836.088
M              0.5158      0.010     50.576      0.000       0.496       0.536
==============================================================================
Omnibus:                       54.536   Durbin-Watson:                   0.695
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              326.201
Skew:                           0.708   Prob(JB):                     1.47e-71
Kurtosis:                       8.533   Cond. No.                     3.66e+04
==============================================================================

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

[2] The condition number is large, 3.66e+04. This might indicate that there are strong multicollinearity or other numerical problems.

[63]:
```
cme.lin_reg(PRIOR_ABSDEPL_STATS_TS, ['M'], 'Depl_Trades')
```

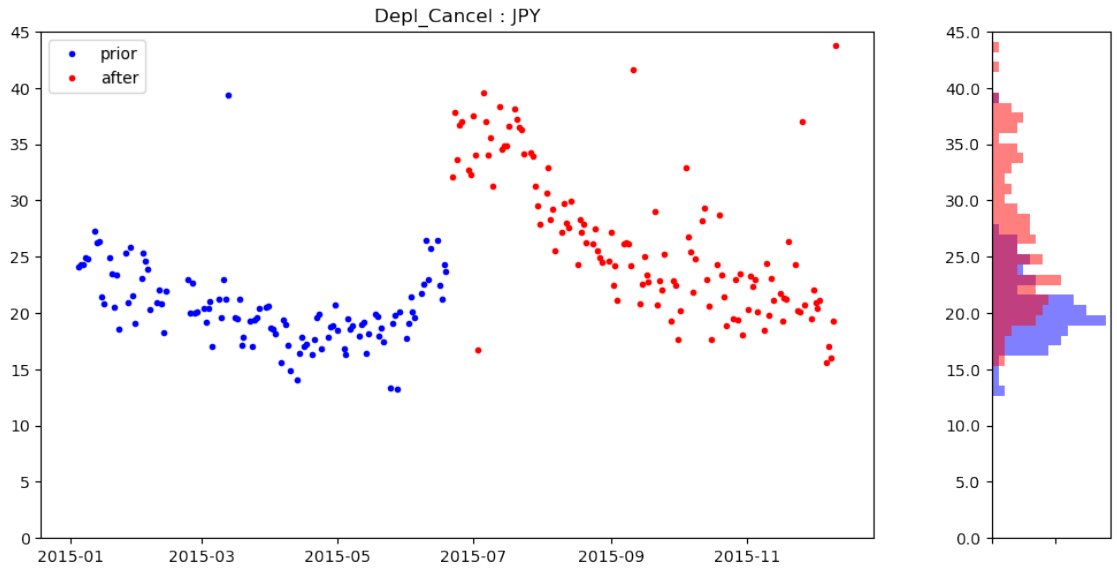
```
                            OLS Regression Results
==============================================================================
Dep. Variable:            Depl_Trades   R-squared:                       0.956
Model:                            OLS   Adj. R-squared:                  0.956
Method:                 Least Squares   F-statistic:                     2462.
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.50e-78
Time:                        16:10:27   Log-Likelihood:                -923.95
No. Observations:                 115   AIC:                             1852.
Df Residuals:                     113   BIC:                             1857.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const      -1613.3969    161.273    -10.004      0.000   -1932.908   -1293.885
M              0.5495      0.011     49.617      0.000       0.528       0.571
==============================================================================
Omnibus:                        4.271   Durbin-Watson:                   1.485
Prob(Omnibus):                  0.118   Jarque-Bera (JB):                4.318
Skew:                           0.241   Prob(JB):                        0.115
Kurtosis:                       3.818   Cond. No.                     3.34e+04
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 3.34e+04. This might indicate that there are strong multicollinearity or other numerical problems.
```

[64]:
```
cme.lin_reg(AFTER_ABSDEPL_STATS_TS, ['M'], 'Depl_Trades')
```

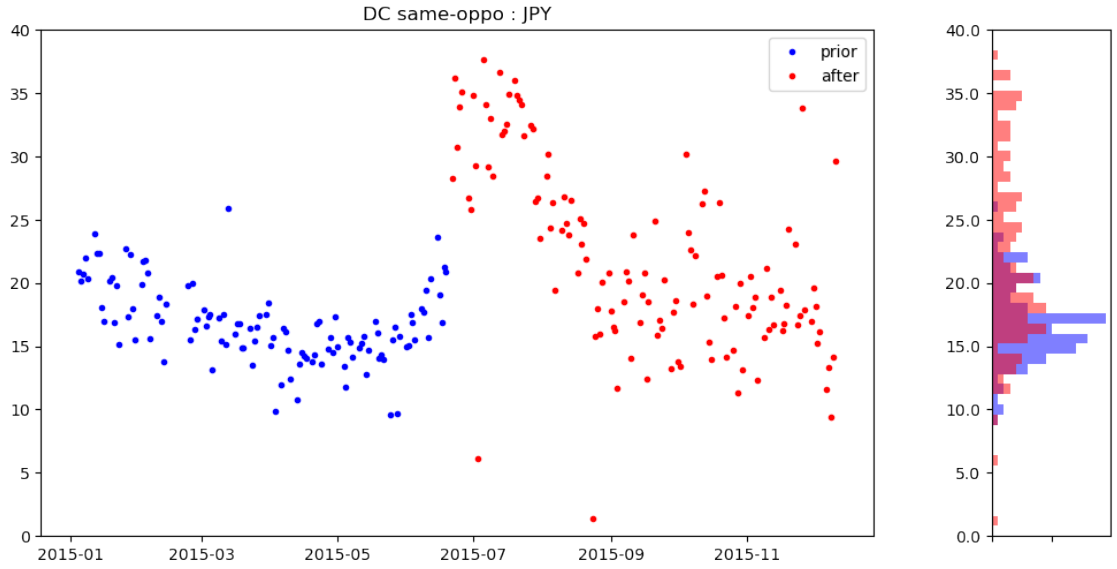
```
                            OLS Regression Results
==============================================================================
Dep. Variable:            Depl_Trades   R-squared:                       0.889
Model:                            OLS   Adj. R-squared:                  0.889
Method:                 Least Squares   F-statistic:                     990.1
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.12e-60
Time:                        16:10:27   Log-Likelihood:                -1101.0
No. Observations:                 125   AIC:                             2206.
Df Residuals:                     123   BIC:                             2212.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
```

```
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       -947.3854    310.675     -3.049      0.003   -1562.347    -332.424
M              0.5015      0.016     31.466      0.000       0.470       0.533
==============================================================================
Omnibus:                       38.024   Durbin-Watson:                   0.596
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              126.173
Skew:                           1.050   Prob(JB):                     4.00e-28
Kurtosis:                       7.452   Cond. No.                     4.15e+04
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 4.15e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

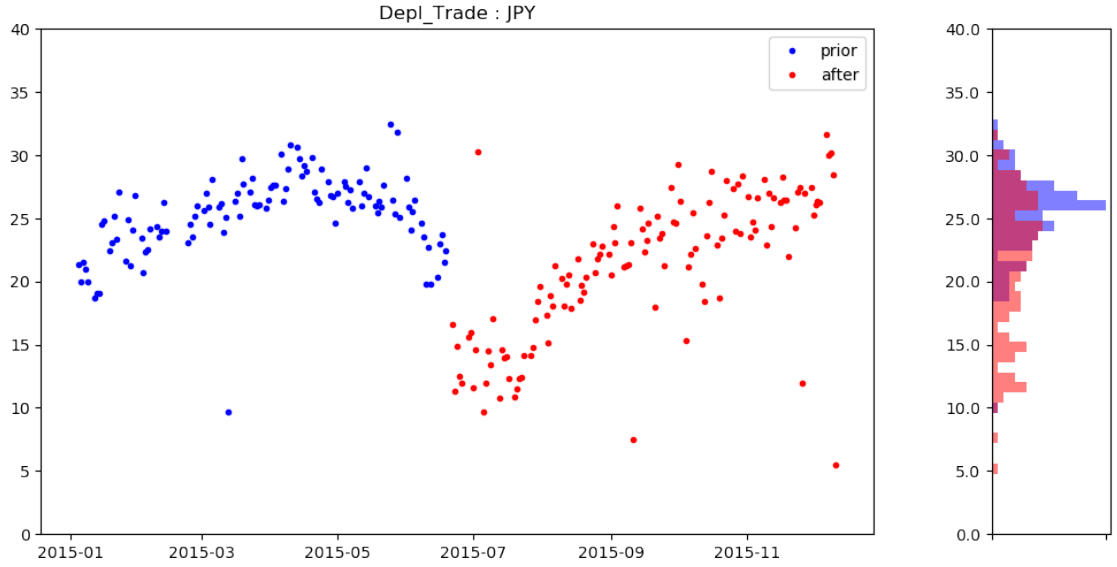
[170]:
```
cme.time_series_hist_plot(DEPL_STATS_TS, 'Depl_Cancel',\
    'Depl_Cancel : '+CURR, 0, 45, 50)
```



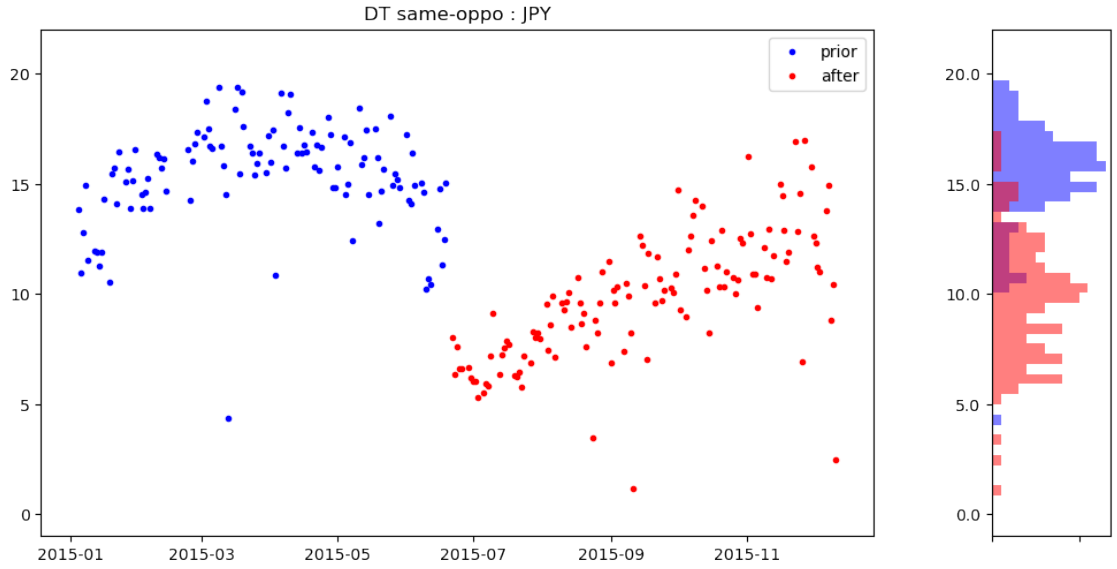
[172]:
```
cme.time_series_hist_plot(DEPL_STATS_TS, 'DC same-oppo',\
    'DC same-oppo : '+CURR, 0, 40, 50)
```



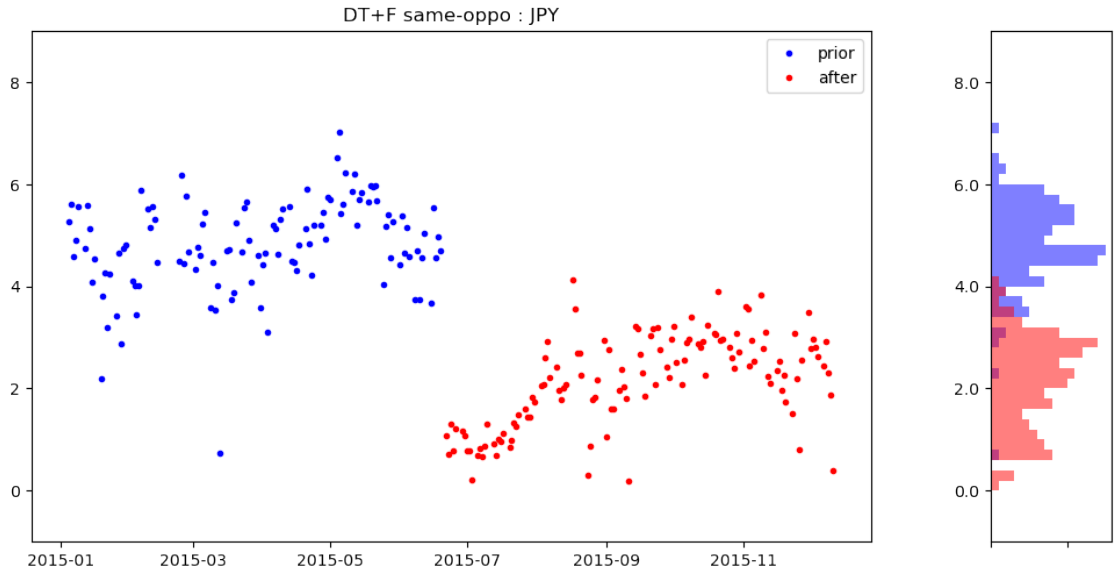
[67]:
```
cme.time_series_hist_plot(DEPL_STATS_TS, 'Depl_Trade',\
    'Depl_Trade : '+CURR, 0, 40, 50)
```



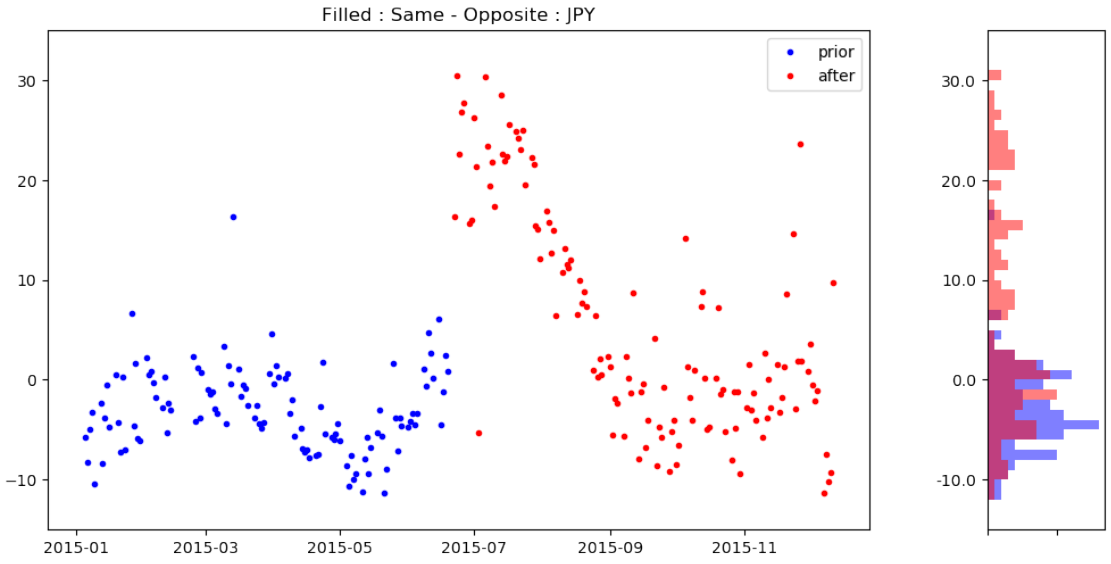
[176]:
```
cme.time_series_hist_plot(DEPL_STATS_TS, 'DT same-oppo',\
    'DT same-oppo : '+CURR, -1, 22, 50)
```



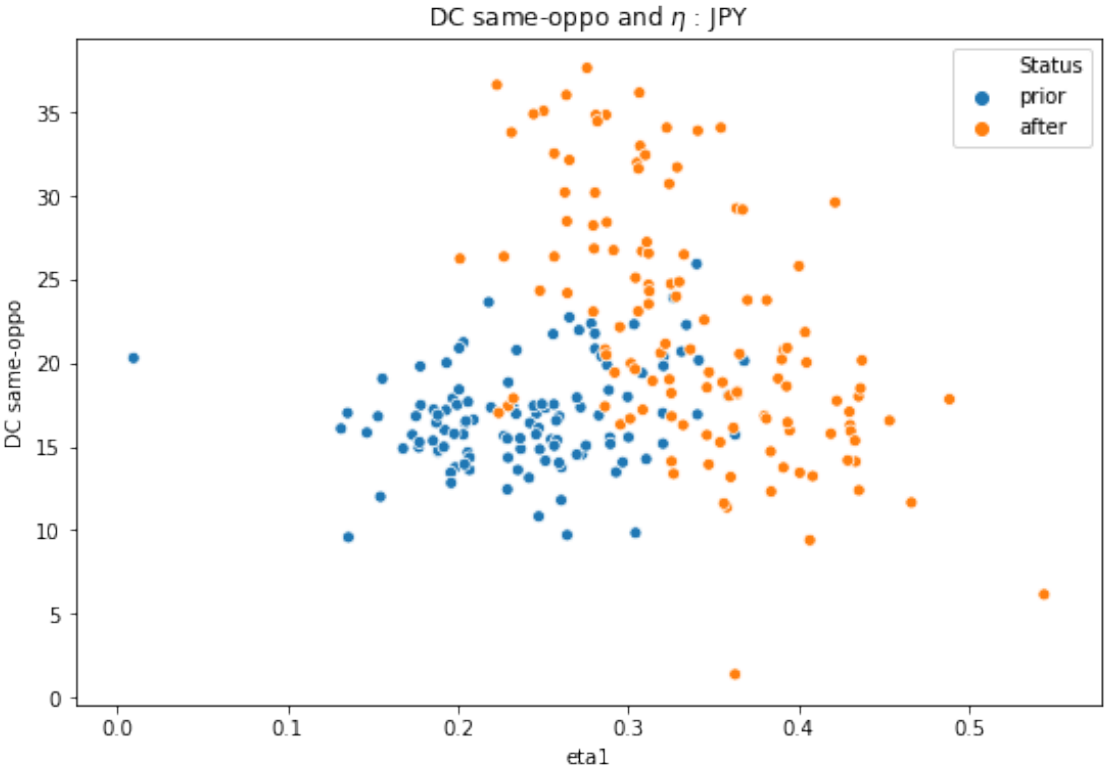
[178]:
```
cme.time_series_hist_plot(DEPL_STATS_TS, 'DT+F same-oppo',\
    'DT+F same-oppo : '+CURR, -1, 9, 50)
```



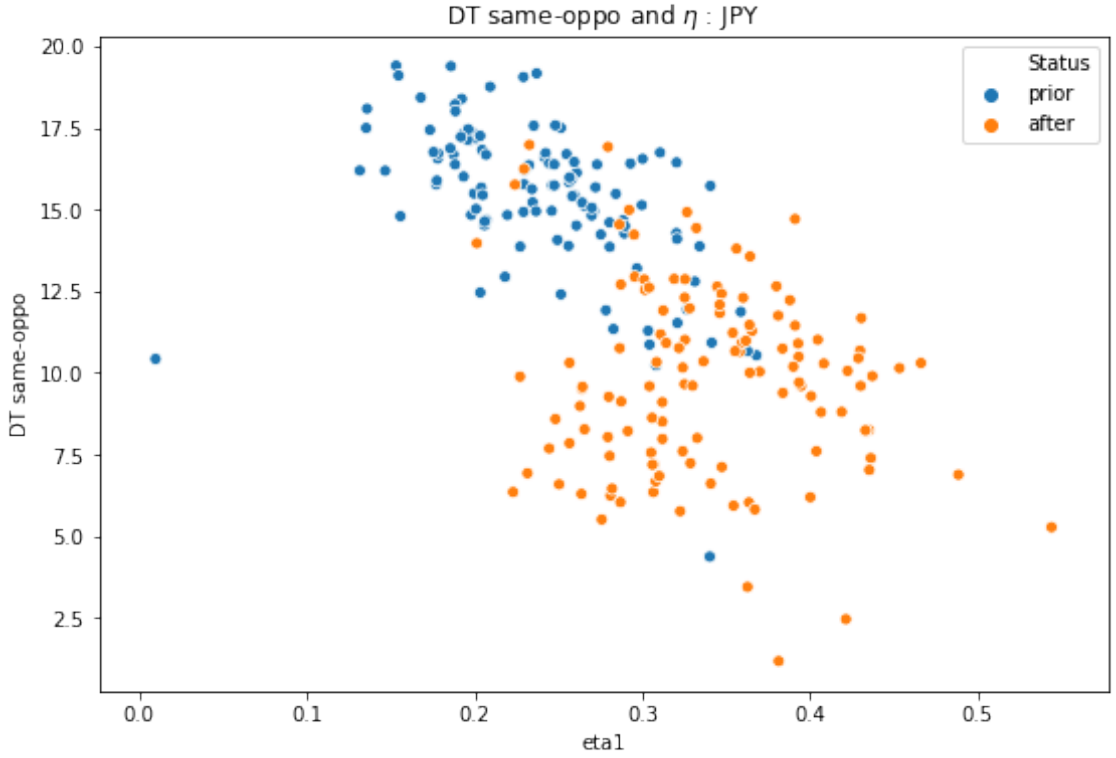
[181]:
```
cme.time_series_hist_plot(DEPL_STATS_TS, 'Fill same-oppo',\
    'Filled : Same - Opposite : '+CURR, -15, 35, 50)
```



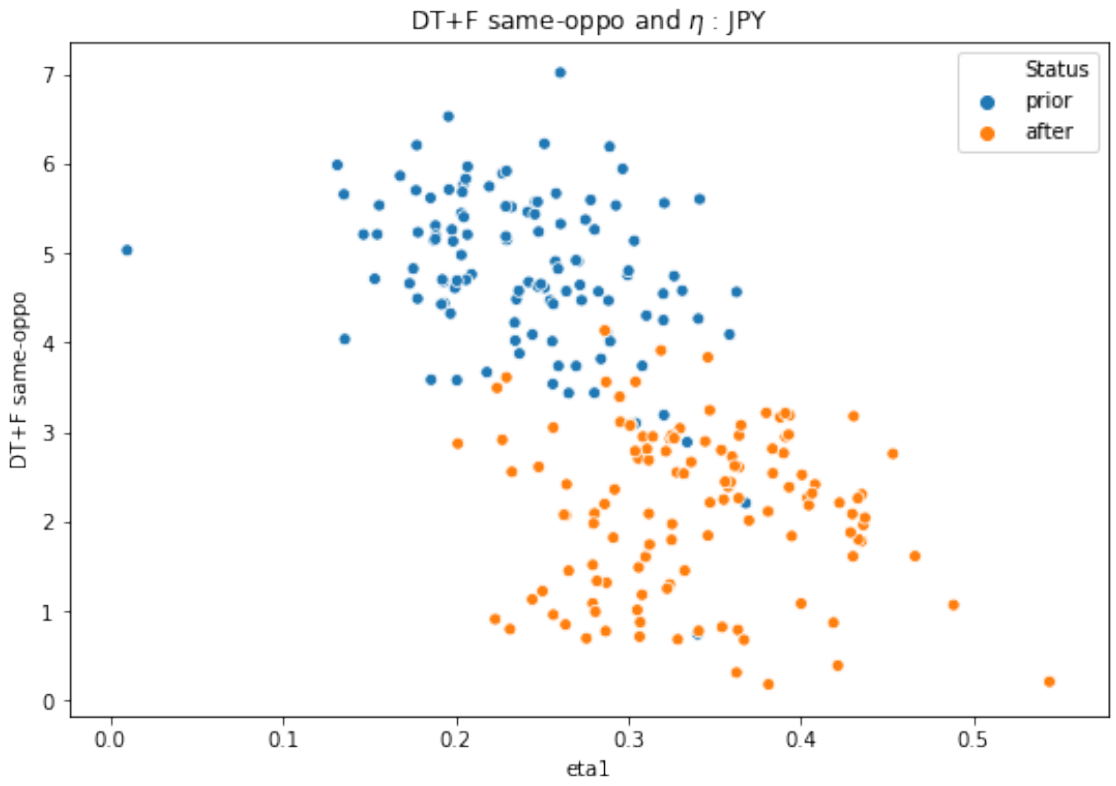
[71]:
```
plt.figure(figsize=(9, 6))
sns.scatterplot(x='eta1', y='DC same-oppo', hue='Status',\
            data=DEPL_STATS_TS);
plt.title('DC same-oppo and $\eta$ : '+CURR);
```



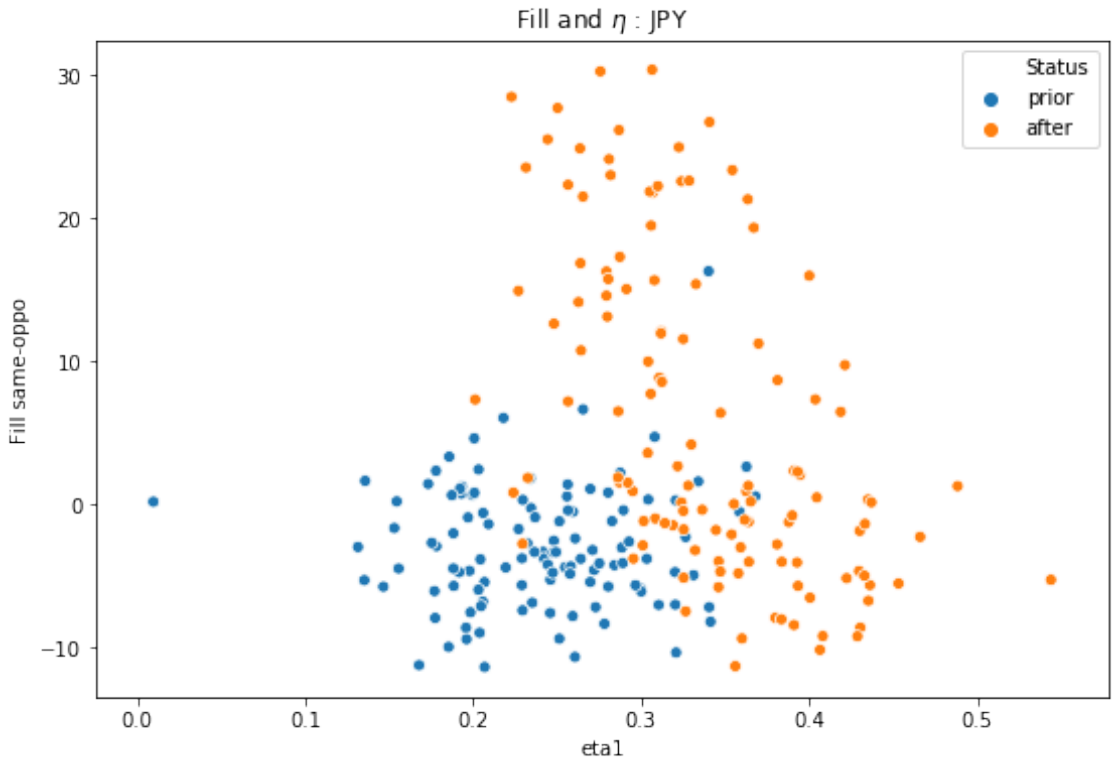
[72]:
```
plt.figure(figsize=(9, 6))
sns.scatterplot(x='eta1', y='DT same-oppo', hue='Status',\
            data=DEPL_STATS_TS);
plt.title('DT same-oppo and $\eta$ : '+CURR);
```



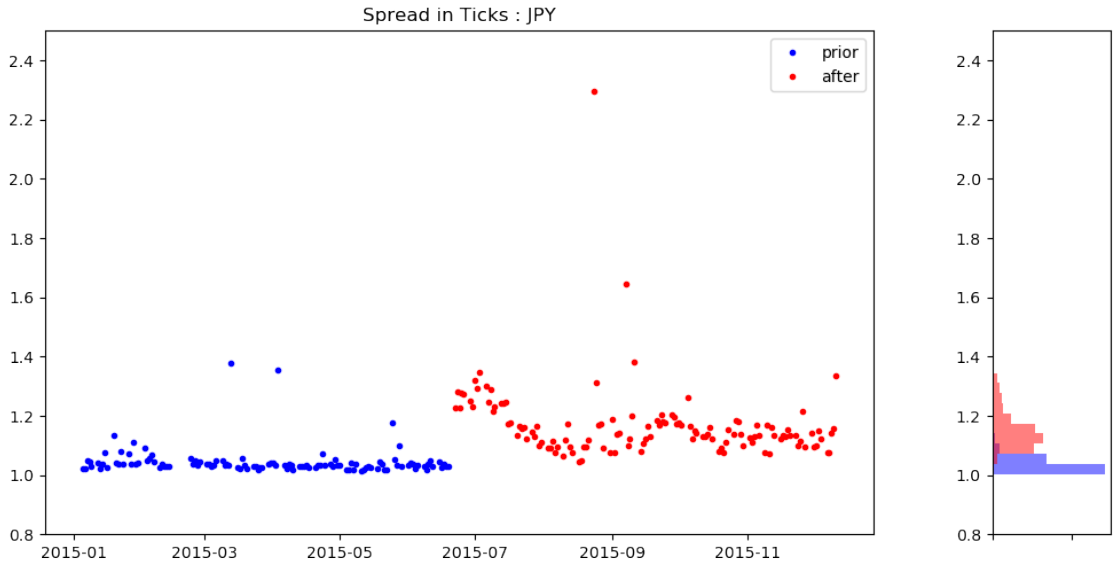
[73]:
```
plt.figure(figsize=(9, 6))
sns.scatterplot(x='eta1', y='DT+F same-oppo', hue='Status',\
            data=DEPL_STATS_TS);
plt.title('DT+F same-oppo and $\eta$ : '+CURR);
```



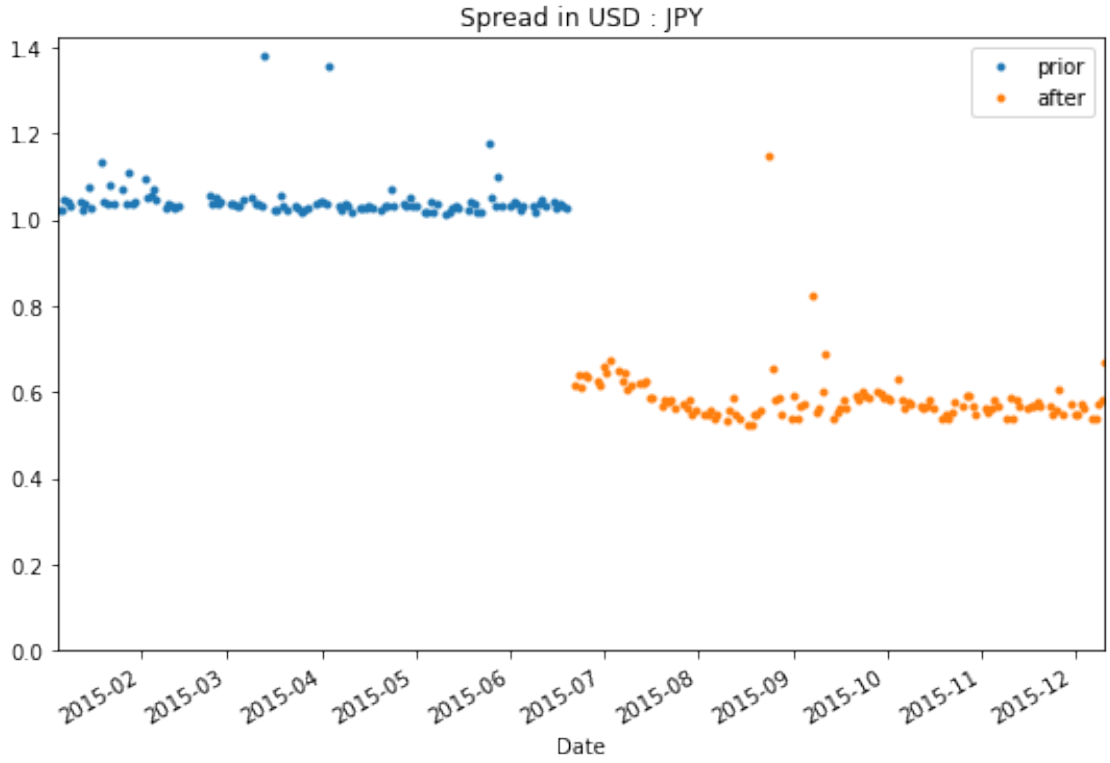
[74]:
```
plt.figure(figsize=(9, 6))
sns.scatterplot(x='eta1', y='Fill same-oppo', hue='Status',\
           data=DEPL_STATS_TS);
plt.title('Fill and $\eta$ : '+CURR);
```



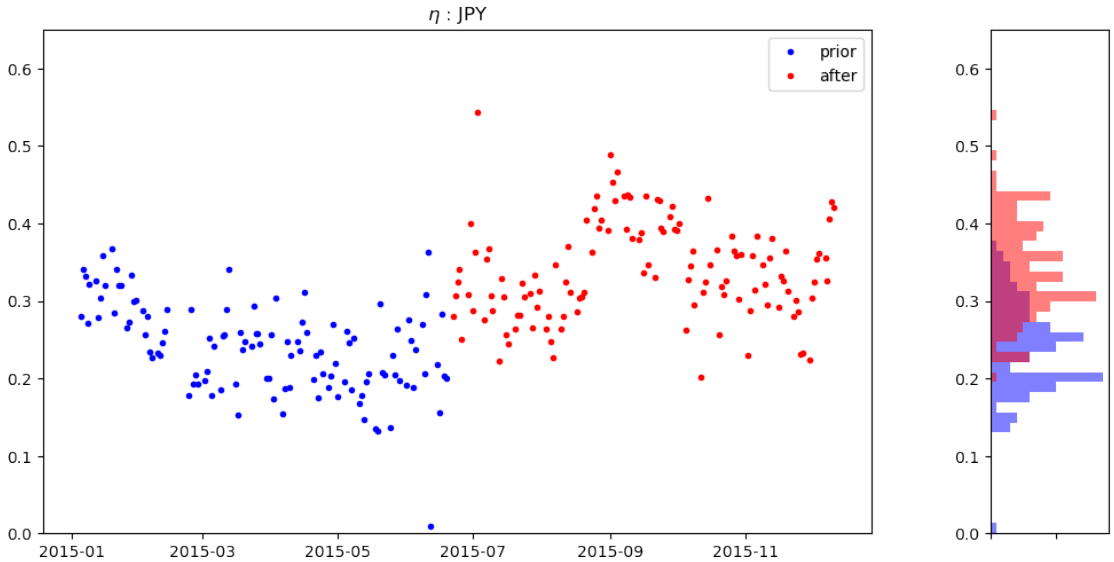
[183]:
```
cme.time_series_hist_plot(OB_UZ_STATS, 'twspr1',\
    'Spread in Ticks : '+CURR, 0.8, 2.5, 50)
```



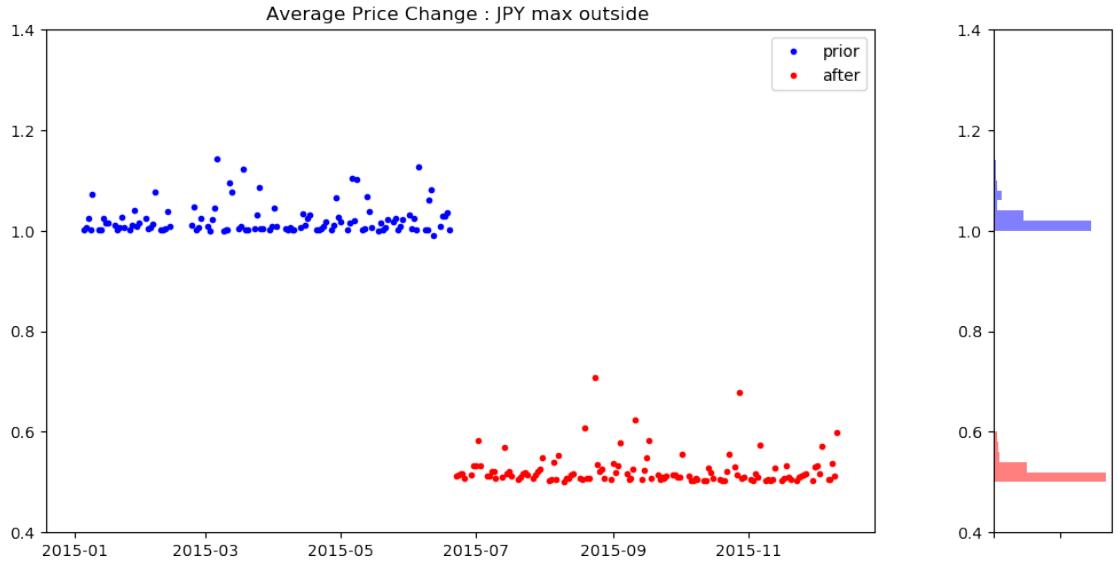
[76]:
```
cme.twspr_plot_USD(OB_UZ_STATS, CURR)
```



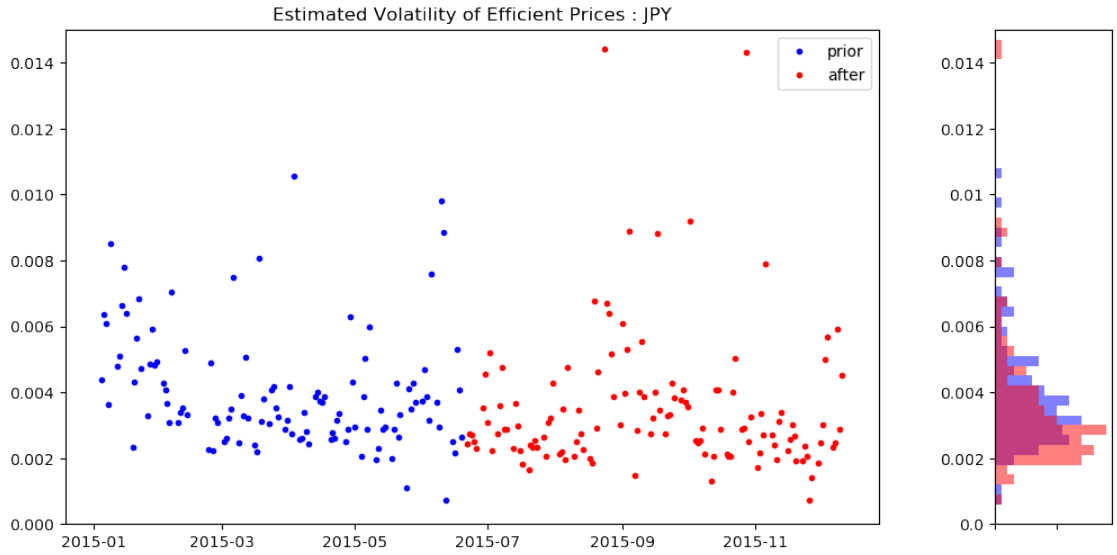
[184]:
```
cme.time_series_hist_plot(OB_UZ_STATS, 'eta1',\
    '$\eta$ : '+CURR, 0, 0.65, 50)
```



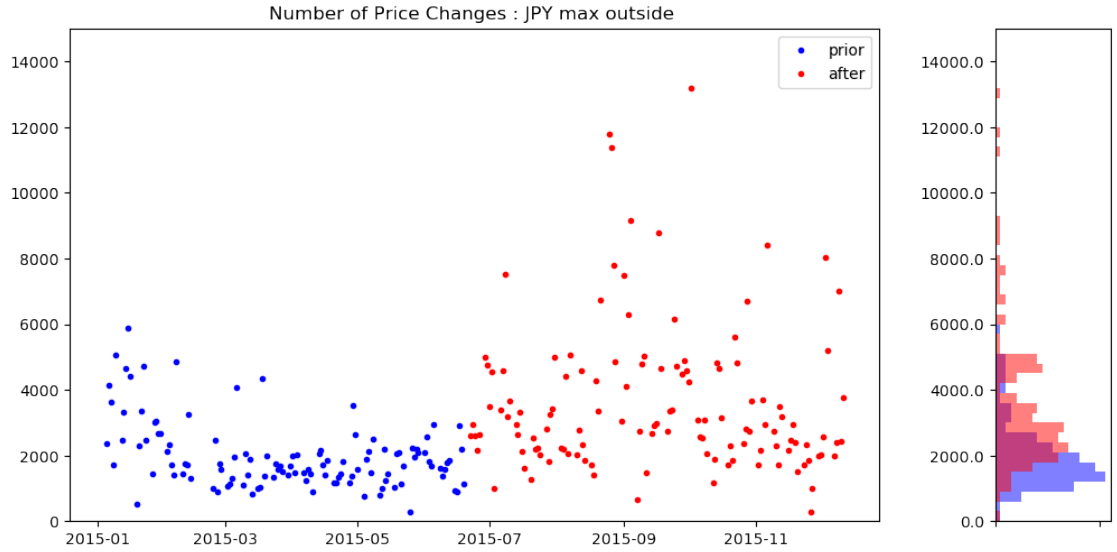
[78]:
```
cme.time_series_hist_plot(OB_UZ_STATS, 'chgavg',\
    'Average Price Change : '+CURR, 0.4, 1.4, 50)
```



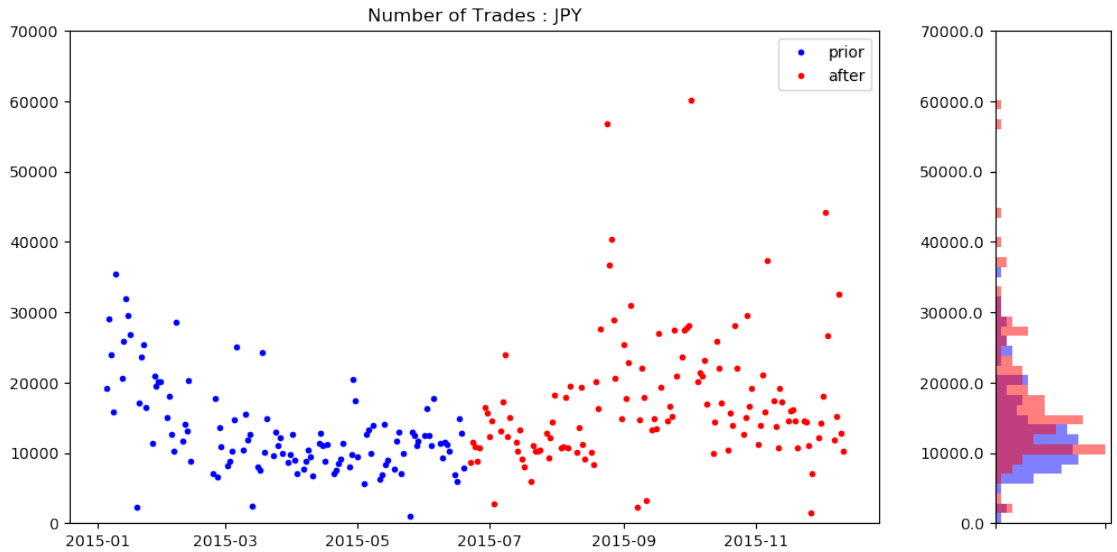
[186]:
```
cme.time_series_hist_plot(OB_UZ_STATS, 'rvxe',\
    'Estimated Volatility of Efficient Prices : '+CURR, 0, 0.015, 50)
```



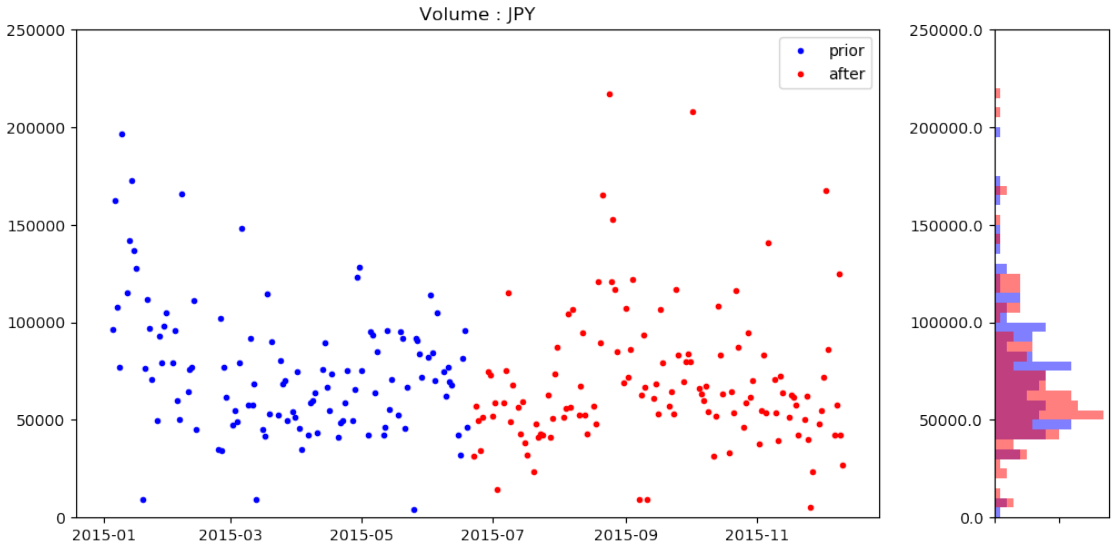
[188]:
```
cme.time_series_hist_plot(OB_UZ_STATS, 'ndfpr',\
    'Number of Price Changes : '+CURR, 0, 15000, 50)
```



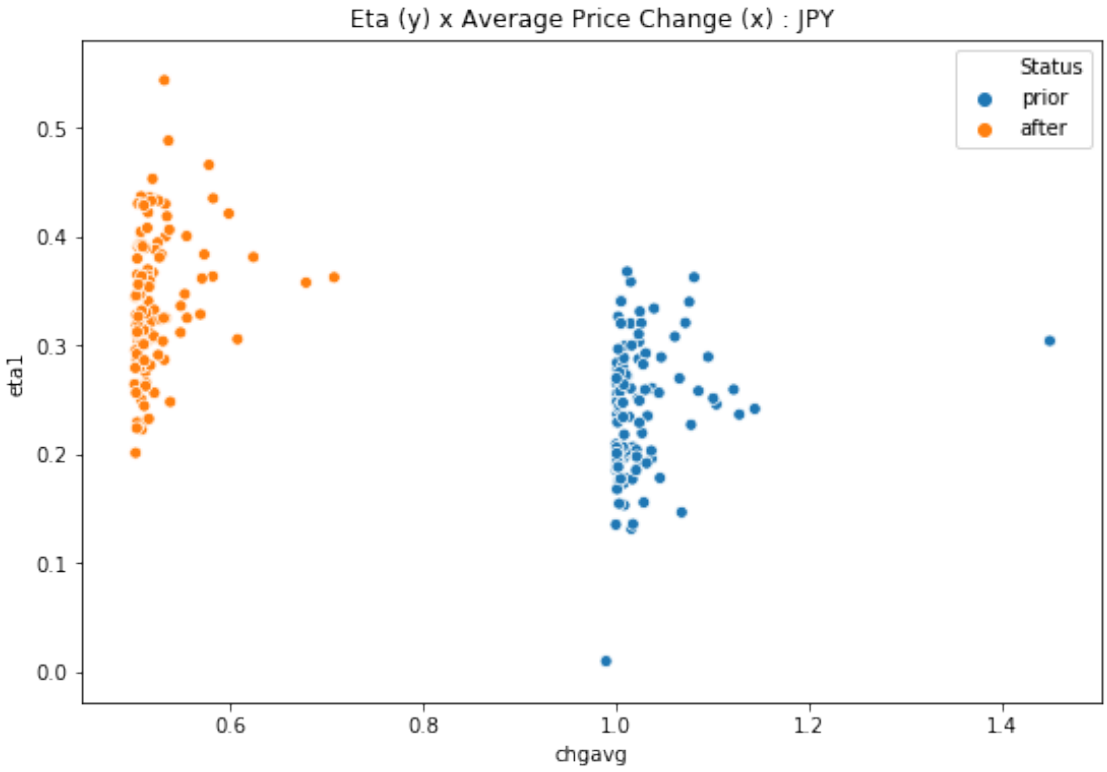
[190]:
```
cme.time_series_hist_plot(OB_UZ_STATS, 'M',\
    'Number of Trades : '+CURR, 0, 70000, 50)
```



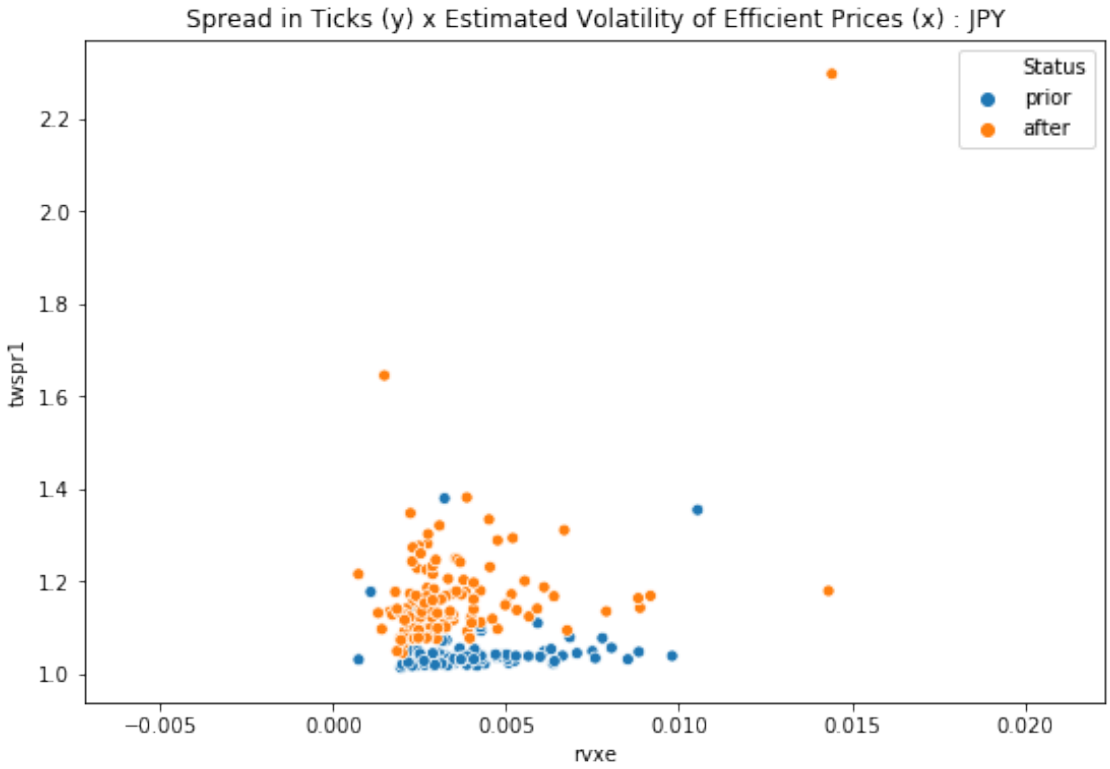
[192]:
```
cme.time_series_hist_plot(OB_UZ_STATS, 'Volume',\
    'Volume : '+CURR, 0, 250000, 50)
```



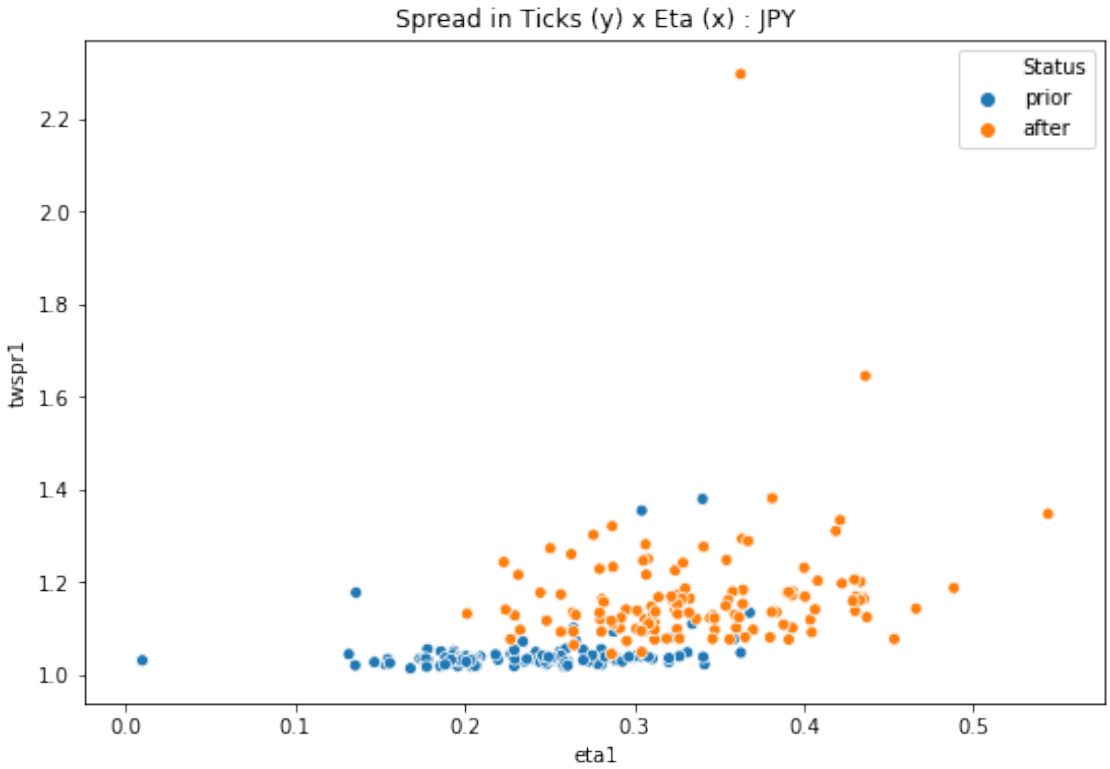
[83]:
```
cme.scatter_plot(OB_UZ_STATS, 'chgavg', 'eta1',\
    'Eta (y) x Average Price Change (x) : '+CURR)
```



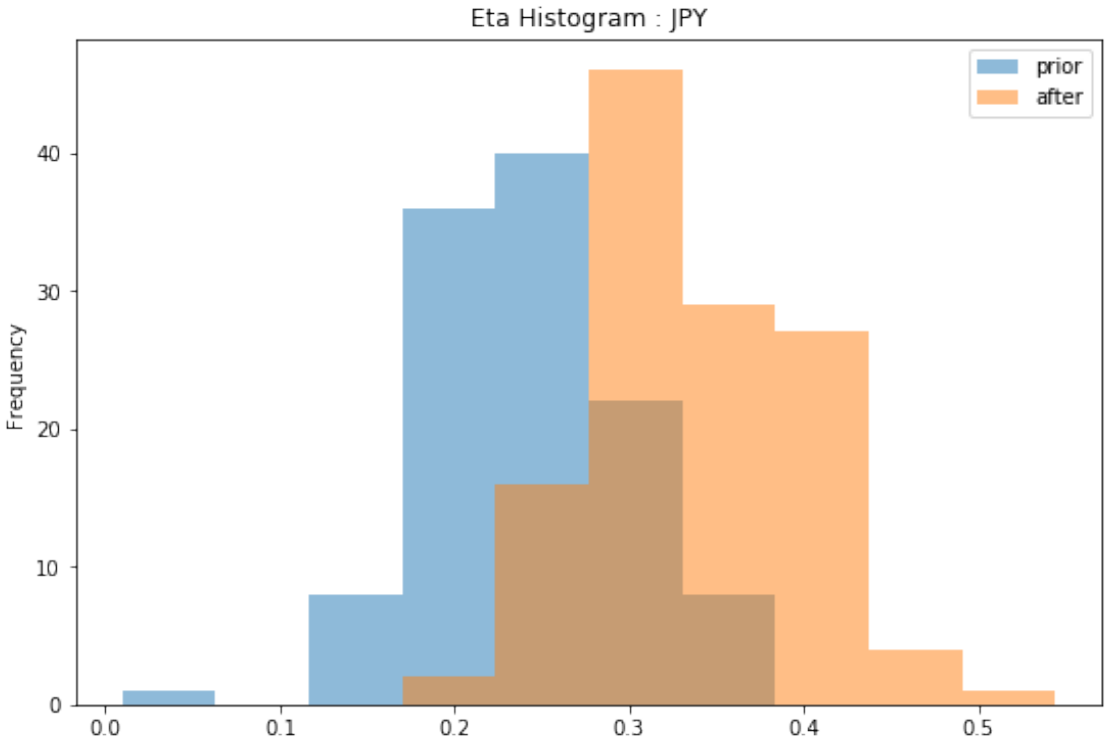
[84]:
```
cme.scatter_plot(OB_UZ_STATS, 'rvxe', 'twspr1',\
    'Spread in Ticks (y) x Estimated Volatility of Efficient Prices (x) :␣
↪'+CURR)
```



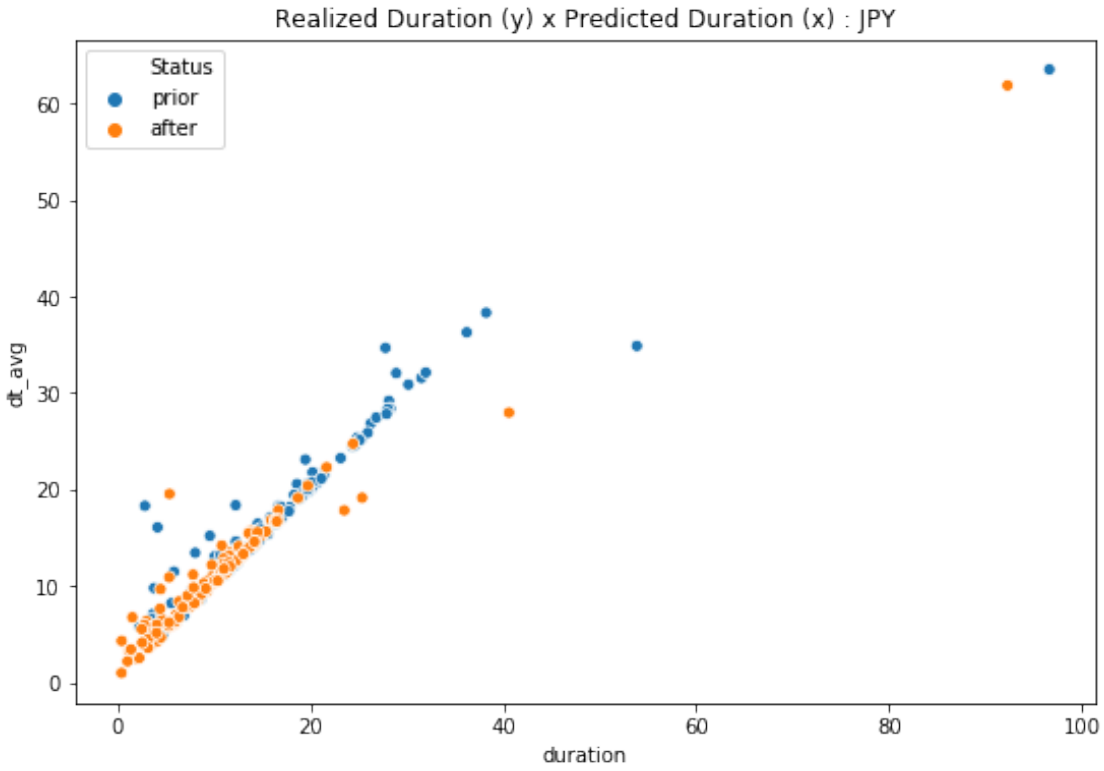
[85]:
```
cme.scatter_plot(OB_UZ_STATS, 'eta1', 'twspr1',\
    'Spread in Ticks (y) x Eta (x) : '+CURR)
```



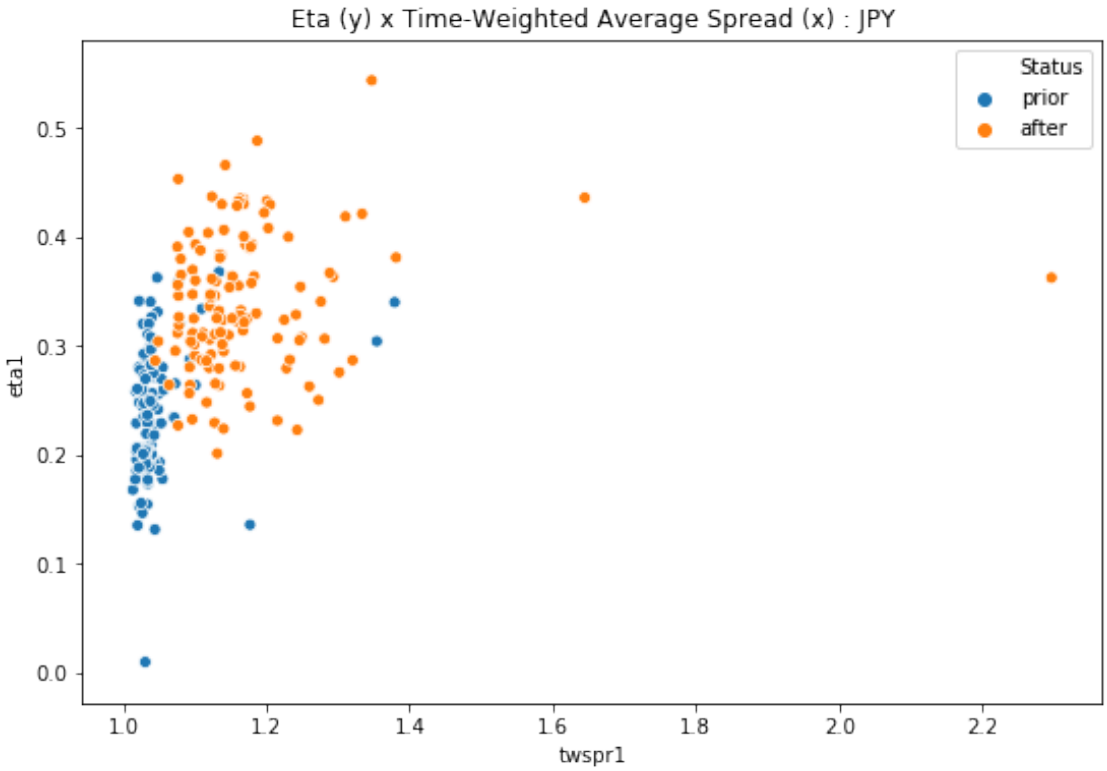
[86]:
```
cme.time_series_hist(OB_UZ_STATS, 'eta1',\
    'Eta Histogram : '+CURR)
```



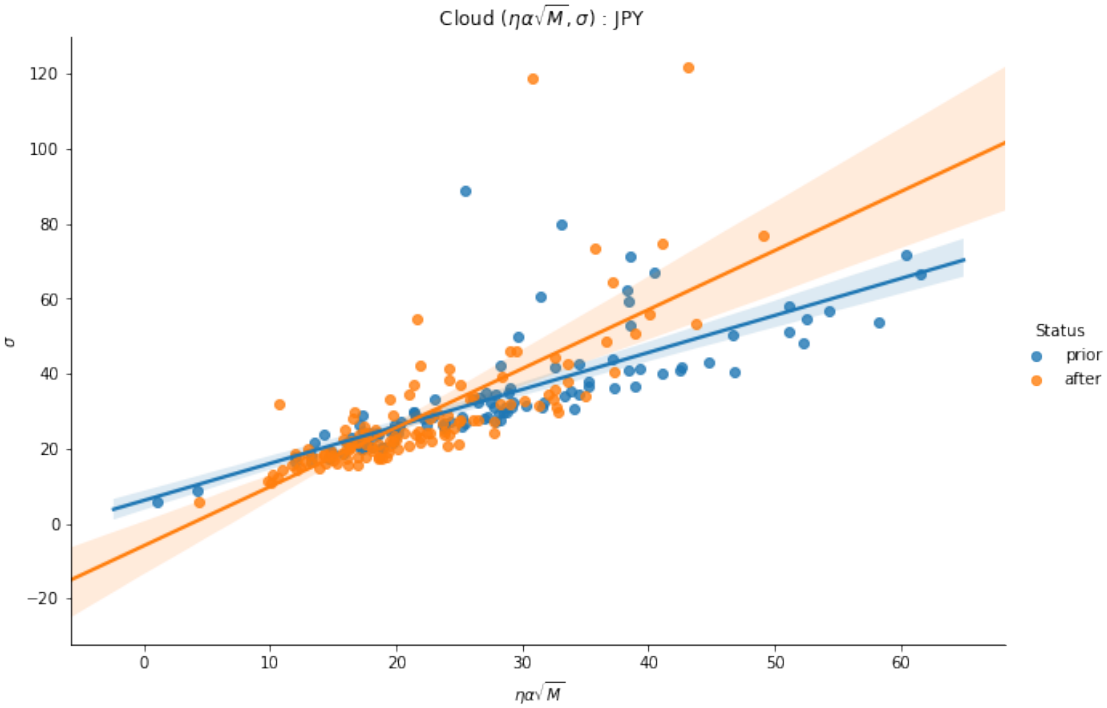
[87]:
```
cme.scatter_plot(OB_UZ_STATS, 'duration', 'dt_avg',\
    'Realized Duration (y) x Predicted Duration (x) : '+CURR)
```



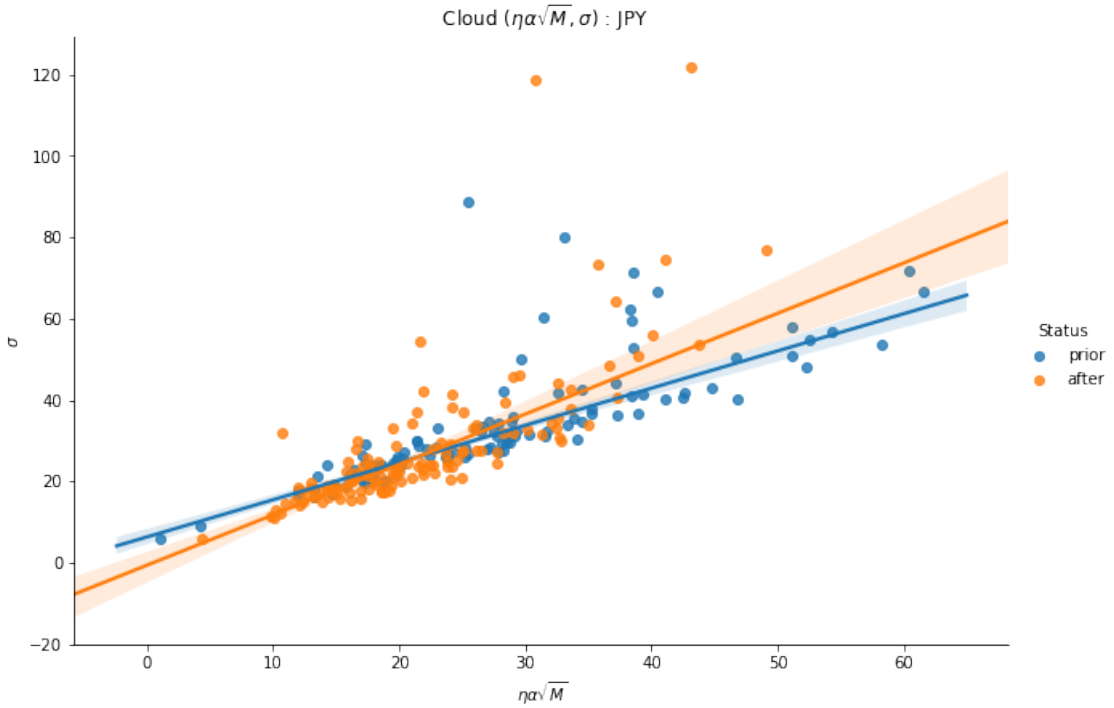
[88]:
```
cme.scatter_plot(OB_UZ_STATS, 'twspr1', 'eta1',\
    'Eta (y) x Time-Weighted Average Spread (x) : '+CURR)
```



[89]:
```
cme.cloud1(OB_UZ_STATS, CURR)
```



[90]:
```
cme.cloud1(OB_UZ_STATS, CURR, True)
```



[91]:
```
cme.lin_reg(PRIOR_OB_UZ_STATS, ['eta*alpha*sqrt(M)', 'S*sqrt(M)'], 'sigma')
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  sigma   R-squared:                       0.629
Model:                            OLS   Adj. R-squared:                  0.623
Method:                 Least Squares   F-statistic:                     95.04
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           7.40e-25
Time:                        16:10:48   Log-Likelihood:                -414.17
No. Observations:                 115   AIC:                             834.3
Df Residuals:                     112   BIC:                             842.6
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.5655      4.570     -0.124      0.902      -9.621
```

```
8.490
eta*alpha*sqrt(M)     0.7324      0.168      4.371      0.000       0.400
1.064
S*sqrt(M)             0.1185      0.071      1.680      0.096      -0.021
0.258
==============================================================================
Omnibus:                      125.596   Durbin-Watson:                   1.696
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1959.511
Skew:                           3.873   Prob(JB):                         0.00
Kurtosis:                      21.680   Cond. No.                         670.
==============================================================================

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

[92]:
```
cme.lin_reg_rob(PRIOR_OB_UZ_STATS, ['eta*alpha*sqrt(M)', 'S*sqrt(M)'], 'sigma')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                  sigma   No. Observations:                  115
Model:                            RLM   Df Residuals:                      112
Method:                          IRLS   Df Model:                            2
Norm:                         HuberT
Scale Est.:                      mad
Cov Type:                         H1
Date:                Wed, 09 Oct 2019
Time:                        16:10:48
No. Iterations:                    50
==============================================================================
=====
                         coef    std err          z      P>|z|      [0.025
0.975]
------------------------------------------------------------------------------
-----
const                 -1.4132      1.786     -0.791      0.429      -4.914
2.087
eta*alpha*sqrt(M)      0.6077      0.065      9.281      0.000       0.479
0.736
S*sqrt(M)              0.1397      0.028      5.063      0.000       0.086
0.194
==============================================================================
=====

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[93]:
```
cme.lin_reg(AFTER_OB_UZ_STATS, ['eta*alpha*sqrt(M)', 'S*sqrt(M)'], 'sigma')
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  sigma   R-squared:                       0.677
Model:                            OLS   Adj. R-squared:                  0.672
Method:                 Least Squares   F-statistic:                     127.8
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.17e-30
Time:                        16:10:48   Log-Likelihood:                -462.49
No. Observations:                 125   AIC:                             931.0
Df Residuals:                     122   BIC:                             939.5
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const        -13.5533      2.791     -4.856      0.000     -19.078      -8.029
eta*alpha*sqrt(M)     0.6932      0.175      3.958      0.000       0.347       1.040
S*sqrt(M)      0.3620      0.057      6.368      0.000       0.250       0.475
==============================================================================
Omnibus:                      140.664   Durbin-Watson:                   1.753
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             3653.972
Skew:                           3.913   Prob(JB):                         0.00
Kurtosis:                      28.304   Cond. No.                         259.
==============================================================================

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

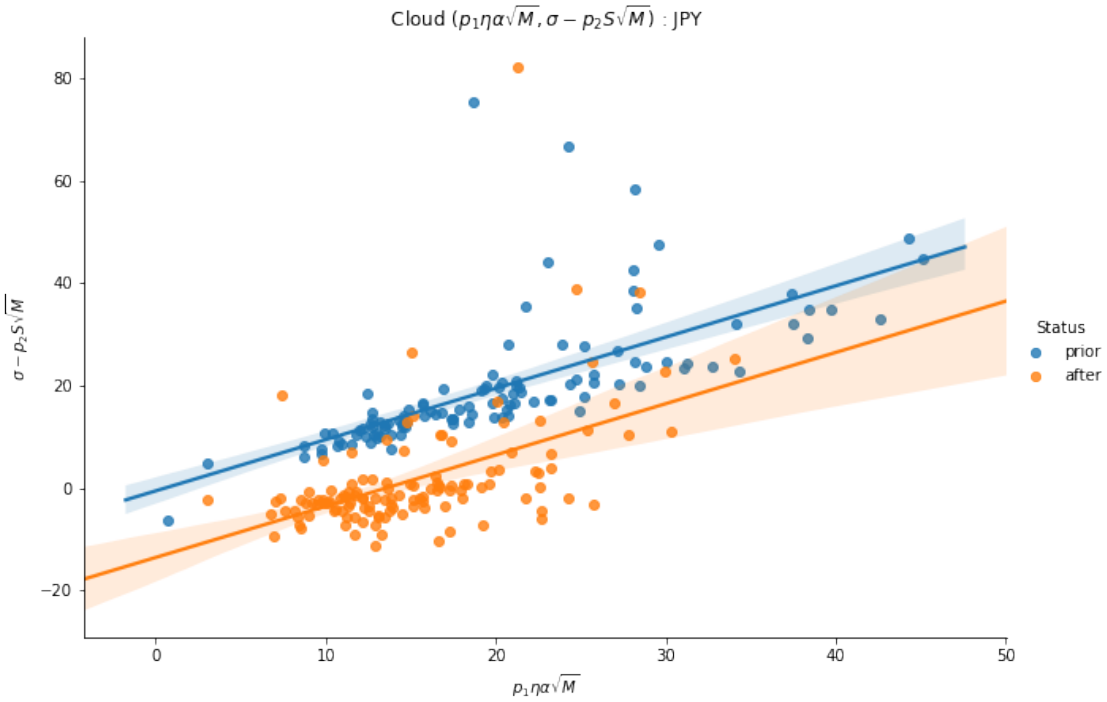
[94]:
```
cme.lin_reg_rob(AFTER_OB_UZ_STATS, ['eta*alpha*sqrt(M)', 'S*sqrt(M)'], 'sigma')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                  sigma   No. Observations:                  125
Model:                            RLM   Df Residuals:                      122
Method:                          IRLS   Df Model:                            2
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:10:48
```

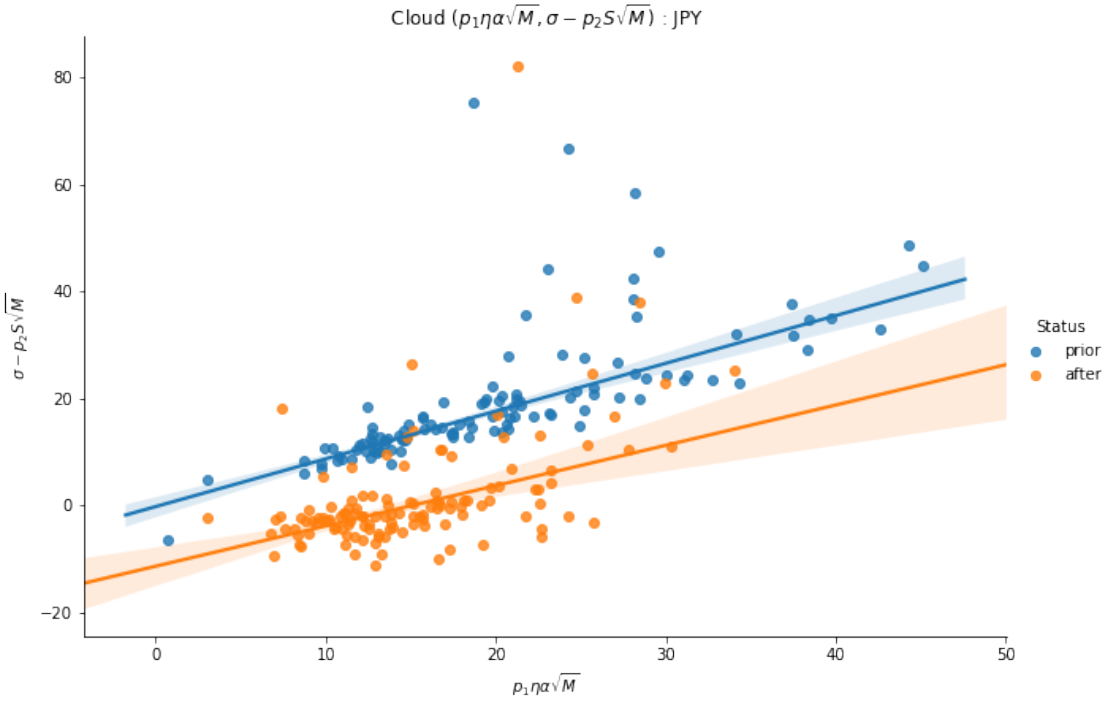
```
No. Iterations:                    50
==============================================================================
=====
                         coef    std err          z      P>|z|      [0.025
0.975]
------------------------------------------------------------------------------
-----
const                -11.4321      1.362     -8.391      0.000     -14.102
-8.762
eta*alpha*sqrt(M)      0.5201      0.085      6.084      0.000       0.353
0.688
S*sqrt(M)              0.3630      0.028     13.080      0.000       0.309
0.417
==============================================================================
=====

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[95]:
```
OB_UZ_STATS['p1*eta*alpha*sqrt(M)'] = np.where(OB_UZ_STATS['Status']=='prior',\
    cme.lin_reg_params(PRIOR_OB_UZ_STATS, ['eta*alpha*sqrt(M)', 'S*sqrt(M)'],␣
 ↪'sigma')['eta*alpha*sqrt(M)'],\
    cme.lin_reg_params(AFTER_OB_UZ_STATS, ['eta*alpha*sqrt(M)', 'S*sqrt(M)'],␣
 ↪'sigma')['eta*alpha*sqrt(M)'])\
    *OB_UZ_STATS['eta*alpha*sqrt(M)']
OB_UZ_STATS['sigma-p2*S*sqrt(M)'] = OB_UZ_STATS['sigma']-\
    np.where(OB_UZ_STATS['Status']=='prior',\
    cme.lin_reg_params(PRIOR_OB_UZ_STATS, ['eta*alpha*sqrt(M)', 'S*sqrt(M)'],␣
 ↪'sigma')['S*sqrt(M)'],\
    cme.lin_reg_params(AFTER_OB_UZ_STATS, ['eta*alpha*sqrt(M)', 'S*sqrt(M)'],␣
 ↪'sigma')['S*sqrt(M)'])*\
    OB_UZ_STATS['S*sqrt(M)']
```

[96]:
```
cme.cloud2(OB_UZ_STATS, CURR)
```



[97]:
```
cme.cloud2(OB_UZ_STATS, CURR, True)
```



[98]:
```
cme.lin_reg(OB_UZ_STATS[OB_UZ_STATS['Status']=='prior'],␣
↪['p1*eta*alpha*sqrt(M)'], 'sigma-p2*S*sqrt(M)')
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     sigma-p2*S*sqrt(M)   R-squared:                       0.480
Model:                            OLS   Adj. R-squared:                  0.475
Method:                 Least Squares   F-statistic:                     104.2
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.00e-17
Time:                        16:11:02   Log-Likelihood:                -414.17
No. Observations:                 115   AIC:                             832.3
Df Residuals:                     113   BIC:                             837.8
Df Model:                           1
Covariance Type:            nonrobust
================================================================================
========
                               coef    std err          t      P>|t|      [0.025
0.975]
--------------------------------------------------------------------------------
--------
const                    -0.5655      2.141     -0.264      0.792      -4.807
3.676
p1*eta*alpha*sqrt(M)      1.0000      0.098     10.205      0.000       0.806
1.194
==============================================================================
Omnibus:                      125.596   Durbin-Watson:                   1.696
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1959.511
Skew:                           3.873   Prob(JB):                         0.00
Kurtosis:                      21.680   Cond. No.                         56.2
==============================================================================

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

[99]:
```
cme.lin_reg_rob(OB_UZ_STATS[OB_UZ_STATS['Status']=='prior'],␣
↪['p1*eta*alpha*sqrt(M)'], 'sigma-p2*S*sqrt(M)')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:     sigma-p2*S*sqrt(M)   No. Observations:                  115
Model:                            RLM   Df Residuals:                      113
Method:                          IRLS   Df Model:                           1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:02
```

```
No. Iterations:                    37
==============================================================================
========
                             coef    std err          z      P>|z|      [0.025
0.975]
------------------------------------------------------------------------------
--------
const                     -0.2559      0.843     -0.304      0.761      -1.908
1.396
p1*eta*alpha*sqrt(M)       0.8932      0.039     23.155      0.000       0.818
0.969
==============================================================================
========

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[100]:
```
cme.lin_reg(OB_UZ_STATS[OB_UZ_STATS['Status']=='after'],␣
↪['p1*eta*alpha*sqrt(M)'], 'sigma-p2*S*sqrt(M)')
```

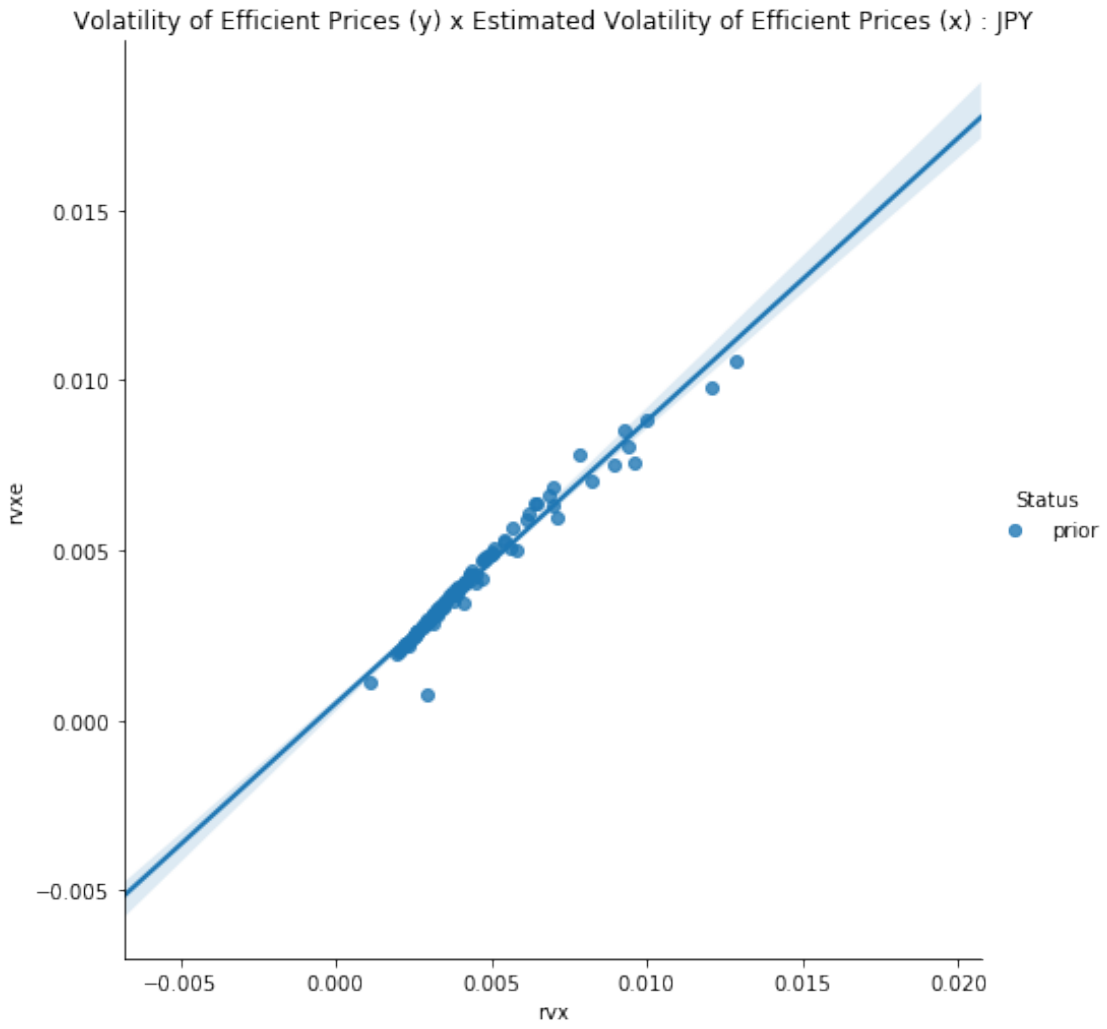
```
                            OLS Regression Results
==============================================================================
Dep. Variable:     sigma-p2*S*sqrt(M)   R-squared:                       0.255
Model:                            OLS   Adj. R-squared:                  0.249
Method:                 Least Squares   F-statistic:                     42.04
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.96e-09
Time:                        16:11:02   Log-Likelihood:                -462.49
No. Observations:                 125   AIC:                             929.0
Df Residuals:                     123   BIC:                             934.6
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
========
                             coef    std err          t      P>|t|      [0.025
0.975]
------------------------------------------------------------------------------
--------
const                    -13.5533      2.506     -5.408      0.000     -18.514
-8.592
p1*eta*alpha*sqrt(M)       1.0000      0.154      6.484      0.000       0.695
1.305
==============================================================================
Omnibus:                      140.664   Durbin-Watson:                   1.753
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             3653.972
Skew:                           3.913   Prob(JB):                         0.00
Kurtosis:                      28.304   Cond. No.                         46.3
==============================================================================
```

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

[101]:
```
cme.lin_reg_rob(OB_UZ_STATS[OB_UZ_STATS['Status']=='after'],␣
↪['p1*eta*alpha*sqrt(M)'], 'sigma-p2*S*sqrt(M)')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:     sigma-p2*S*sqrt(M)   No. Observations:                  125
Model:                            RLM   Df Residuals:                      123
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:02
No. Iterations:                    50
==============================================================================
========
                             coef    std err          z      P>|z|      [0.025
0.975]
------------------------------------------------------------------------------
--------
const                    -11.4088      1.226     -9.306      0.000     -13.812
-9.006
p1*eta*alpha*sqrt(M)       0.7537      0.075      9.991      0.000       0.606
0.902
==============================================================================
========

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[102]:
```
cme.regr_plot(PRIOR_OB_UZ_STATS, 'rvx', 'rvxe',\
    'Volatility of Efficient Prices (y) x Estimated Volatility of Efficient␣
↪Prices (x) : '+CURR)
```

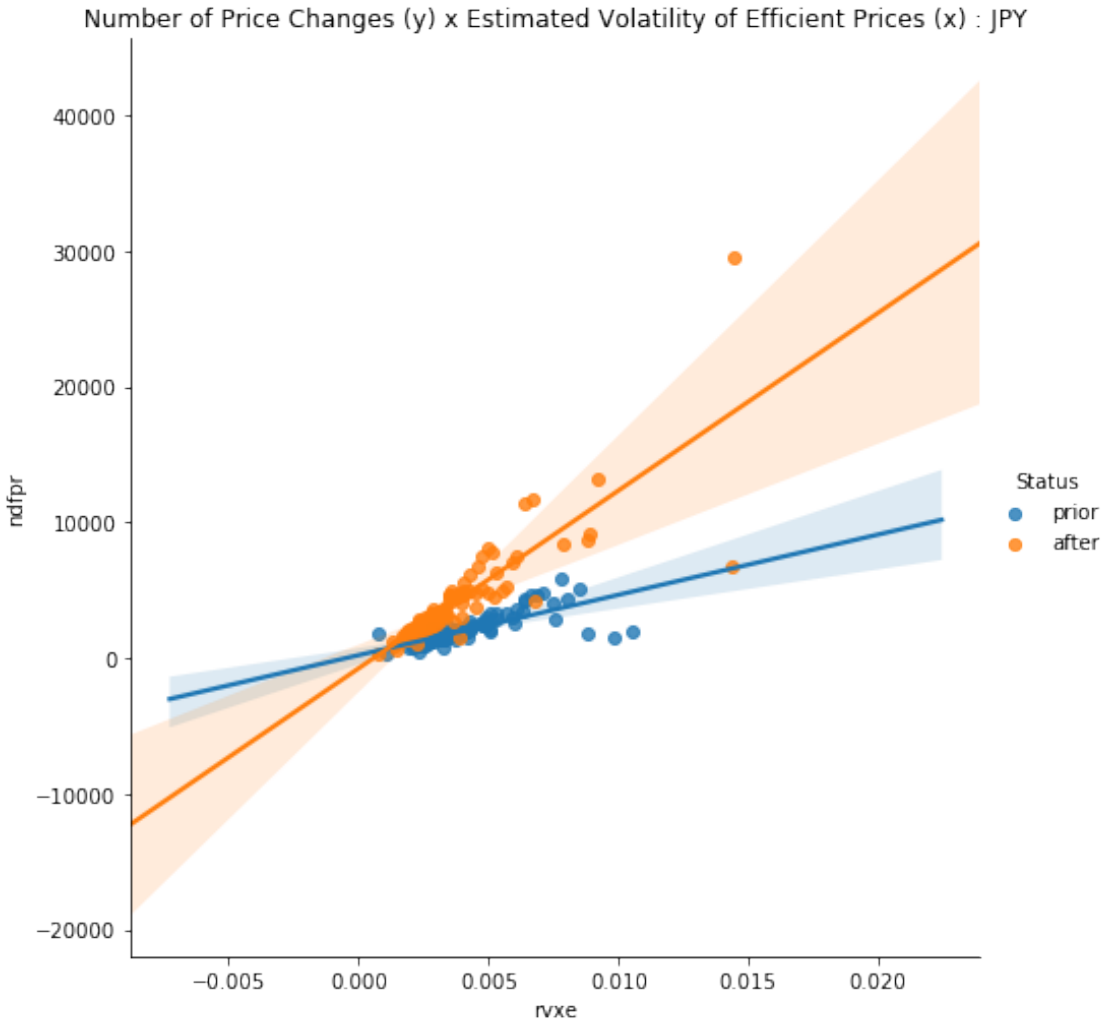


[103]:
```
cme.lin_reg(PRIOR_OB_UZ_STATS, 'rvx', 'rvxe', True)
```

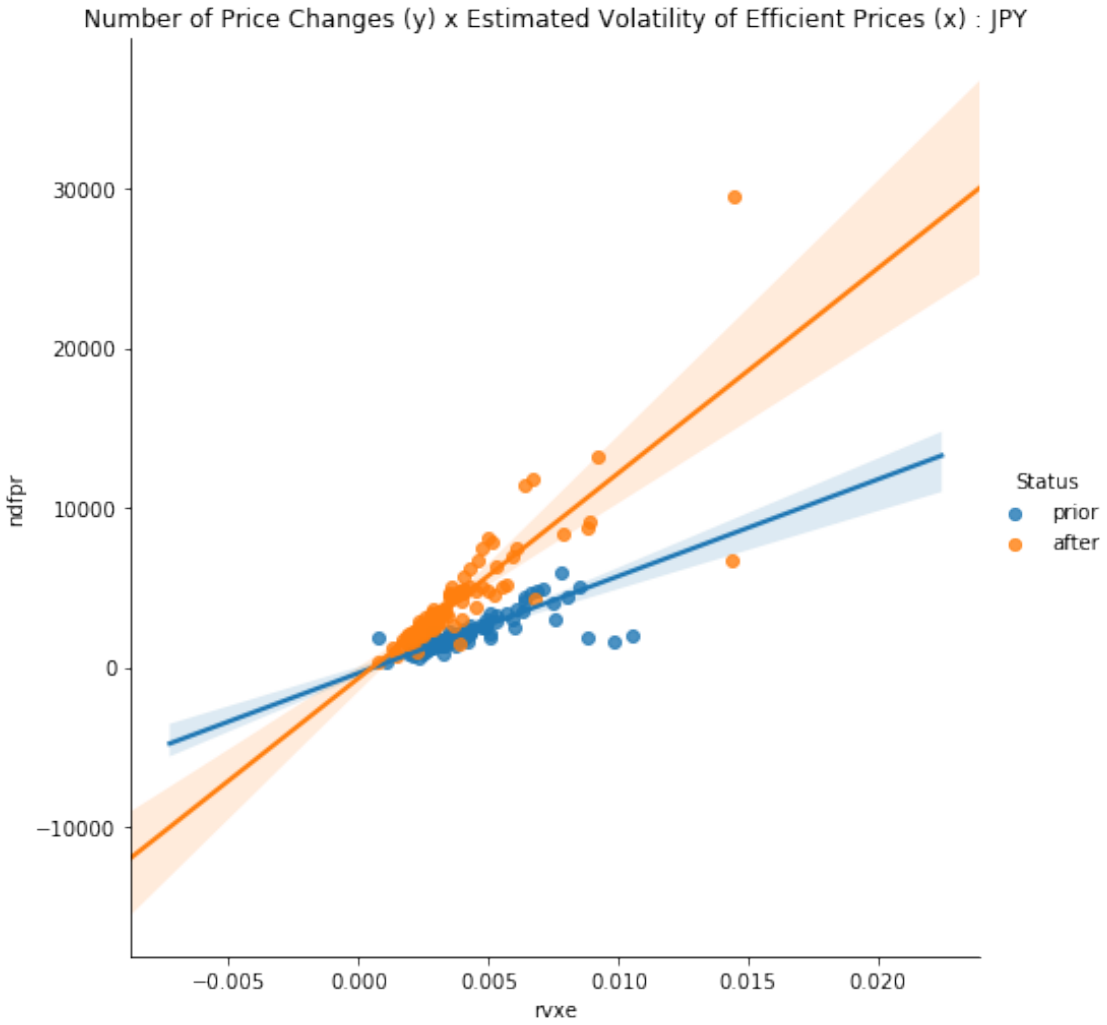
```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   rvxe   R-squared:                       0.897
Model:                            OLS   Adj. R-squared:                  0.896
Method:                 Least Squares   F-statistic:                     981.3
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.53e-57
Time:                        16:11:03   Log-Likelihood:                 69.473
No. Observations:                 115   AIC:                            -134.9
Df Residuals:                     113   BIC:                            -129.5
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

```
const          -0.3790      0.167     -2.265      0.025      -0.711      -0.047
rvx             0.9396      0.030     31.326      0.000       0.880       0.999
==============================================================================
Omnibus:                      230.242   Durbin-Watson:                   1.934
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            40562.325
Skew:                          -9.171   Prob(JB):                         0.00
Kurtosis:                      93.160   Cond. No.                         77.5
==============================================================================

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

[104]:
```
cme.regr_plot(OB_UZ_STATS, 'rvxe', 'ndfpr',\
    'Number of Price Changes (y) x Estimated Volatility of Efficient Prices (x)␣
 ↪: '+CURR)
```



[105]:
```
cme.regr_plot(OB_UZ_STATS, 'rvxe', 'ndfpr',\
    'Number of Price Changes (y) x Estimated Volatility of Efficient Prices (x)␣
↪: '+CURR, True)
```



[106]:
```
cme.lin_reg(PRIOR_OB_UZ_STATS, 'rvxe', 'ndfpr', True)
```

OLS Regression Results

Dep. Variable:	ndfpr	R-squared:	0.640
Model:	OLS	Adj. R-squared:	0.636
Method:	Least Squares	F-statistic:	200.5
Date:	Wed, 09 Oct 2019	Prob (F-statistic):	8.50e-27
Time:	16:11:16	Log-Likelihood:	-20.584
No. Observations:	115	AIC:	45.17

```
Df Residuals:                      113   BIC:                             50.66
Df Model:                            1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         12.7349      0.372     34.241      0.000      11.998      13.472
rvxe           0.9367      0.066     14.161      0.000       0.806       1.068
==============================================================================
Omnibus:                       30.409   Durbin-Watson:                   2.081
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              326.946
Skew:                           0.258   Prob(JB):                     1.01e-71
Kurtosis:                      11.244   Cond. No.                         79.2
==============================================================================

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

[107]:
```
cme.lin_reg_rob(PRIOR_OB_UZ_STATS, 'rvxe', 'ndfpr', True)
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                  ndfpr   No. Observations:                  115
Model:                            RLM   Df Residuals:                      113
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:16
No. Iterations:                    26
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         14.4039      0.178     80.836      0.000      14.055      14.753
rvxe           1.2298      0.032     38.805      0.000       1.168       1.292
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[108]:
```
cme.lin_reg(AFTER_OB_UZ_STATS, 'rvxe', 'ndfpr', True)
```

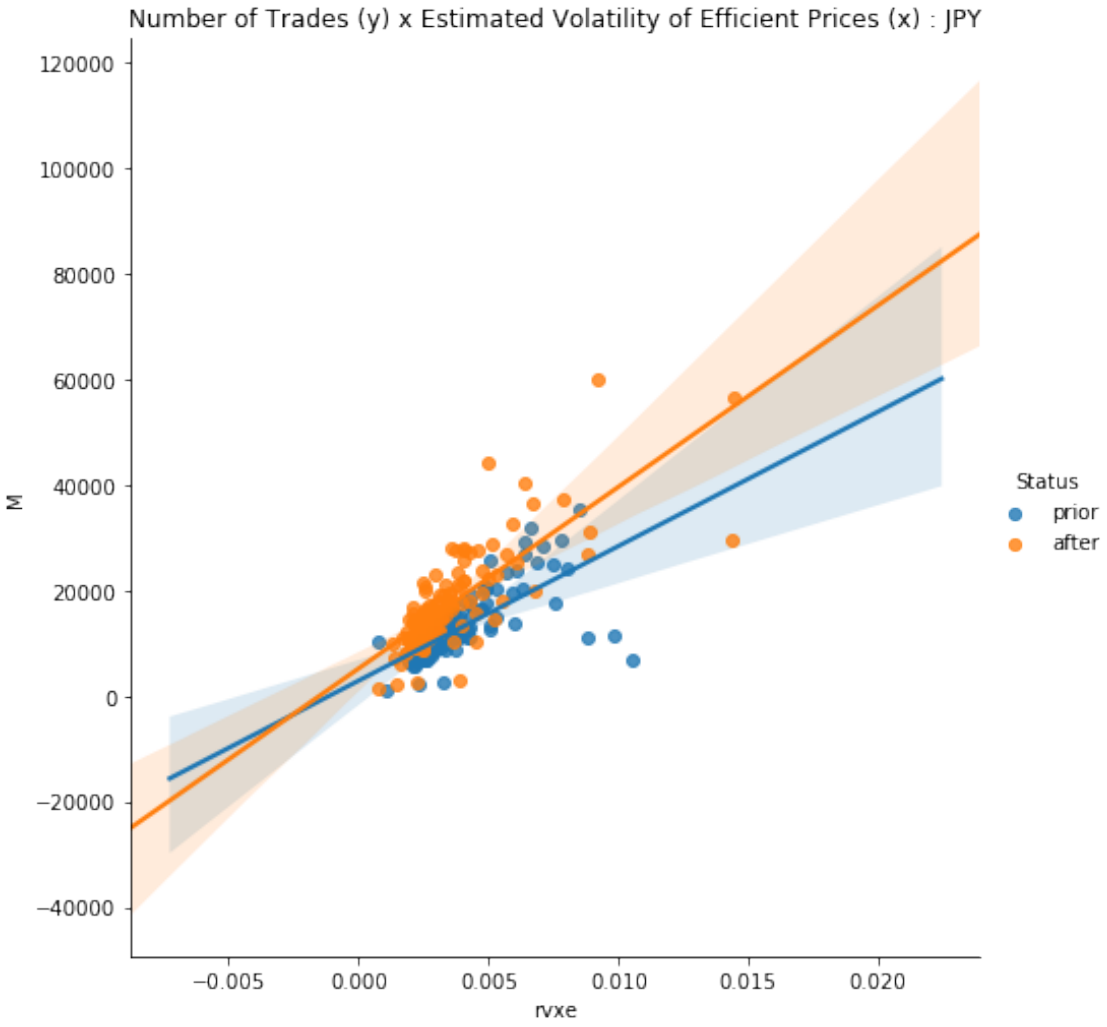
```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  ndfpr   R-squared:                       0.841
Model:                            OLS   Adj. R-squared:                  0.840
```

```
Method:                 Least Squares   F-statistic:                     649.7
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           6.46e-51
Time:                        16:11:16   Log-Likelihood:                0.55552
No. Observations:                 125   AIC:                             2.889
Df Residuals:                     123   BIC:                             8.546
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         15.1309      0.279     54.283      0.000      14.579      15.683
rvxe           1.2265      0.048     25.489      0.000       1.131       1.322
==============================================================================
Omnibus:                       65.568   Durbin-Watson:                   1.618
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              251.141
Skew:                          -1.909   Prob(JB):                     2.92e-55
Kurtosis:                       8.800   Cond. No.                         76.5
==============================================================================

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

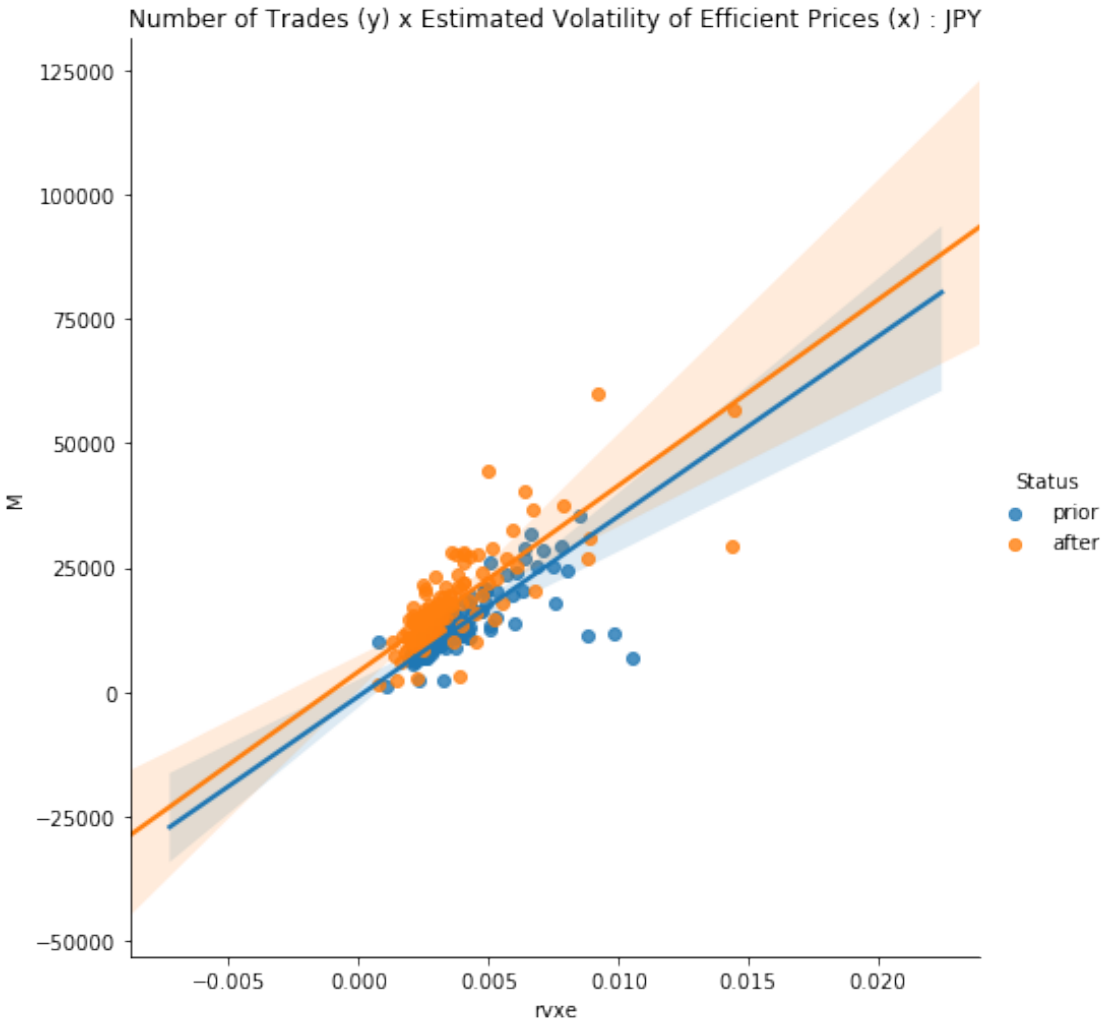
[109]:
```
cme.lin_reg_rob(AFTER_OB_UZ_STATS, 'rvxe', 'ndfpr', True)
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                  ndfpr   No. Observations:                  125
Model:                            RLM   Df Residuals:                      123
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:16
No. Iterations:                    37
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         15.4635      0.203     76.033      0.000      15.065      15.862
rvxe           1.2784      0.035     36.411      0.000       1.210       1.347
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[110]:
```
cme.regr_plot(OB_UZ_STATS, 'rvxe', 'M',\
    'Number of Trades (y) x Estimated Volatility of Efficient Prices (x) :␣
↪'+CURR)
```



[111]:
```
cme.regr_plot(OB_UZ_STATS, 'rvxe', 'M',\
    'Number of Trades (y) x Estimated Volatility of Efficient Prices (x) :␣
↪'+CURR, True)
```



[112]:
```
cme.lin_reg(PRIOR_OB_UZ_STATS, 'rvxe', 'M', True)
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      M   R-squared:                       0.522
Model:                            OLS   Adj. R-squared:                  0.518
Method:                 Least Squares   F-statistic:                     123.4
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           7.87e-20
Time:                        16:11:24   Log-Likelihood:                -44.537
No. Observations:                 115   AIC:                             93.07
Df Residuals:                     113   BIC:                             98.56
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

```
const         14.4390      0.458     31.523      0.000      13.532      15.346
rvxe           0.9051      0.081     11.110      0.000       0.744       1.067
==============================================================================
Omnibus:                       62.069   Durbin-Watson:                   1.838
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              336.094
Skew:                          -1.739   Prob(JB):                    1.04e-73
Kurtosis:                      10.619   Cond. No.                        79.2
==============================================================================

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

[113]:
```
cme.lin_reg_rob(PRIOR_OB_UZ_STATS, 'rvxe', 'M', True)
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                      M   No. Observations:                  115
Model:                            RLM   Df Residuals:                      113
Method:                          IRLS   Df Model:                            1
Norm:                         HuberT
Scale Est.:                      mad
Cov Type:                         H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:24
No. Iterations:                    18
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         15.5520      0.238     65.373      0.000      15.086      16.018
rvxe           1.0956      0.042     25.893      0.000       1.013       1.179
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[114]:
```
cme.lin_reg(AFTER_OB_UZ_STATS, 'rvxe', 'M', True)
```

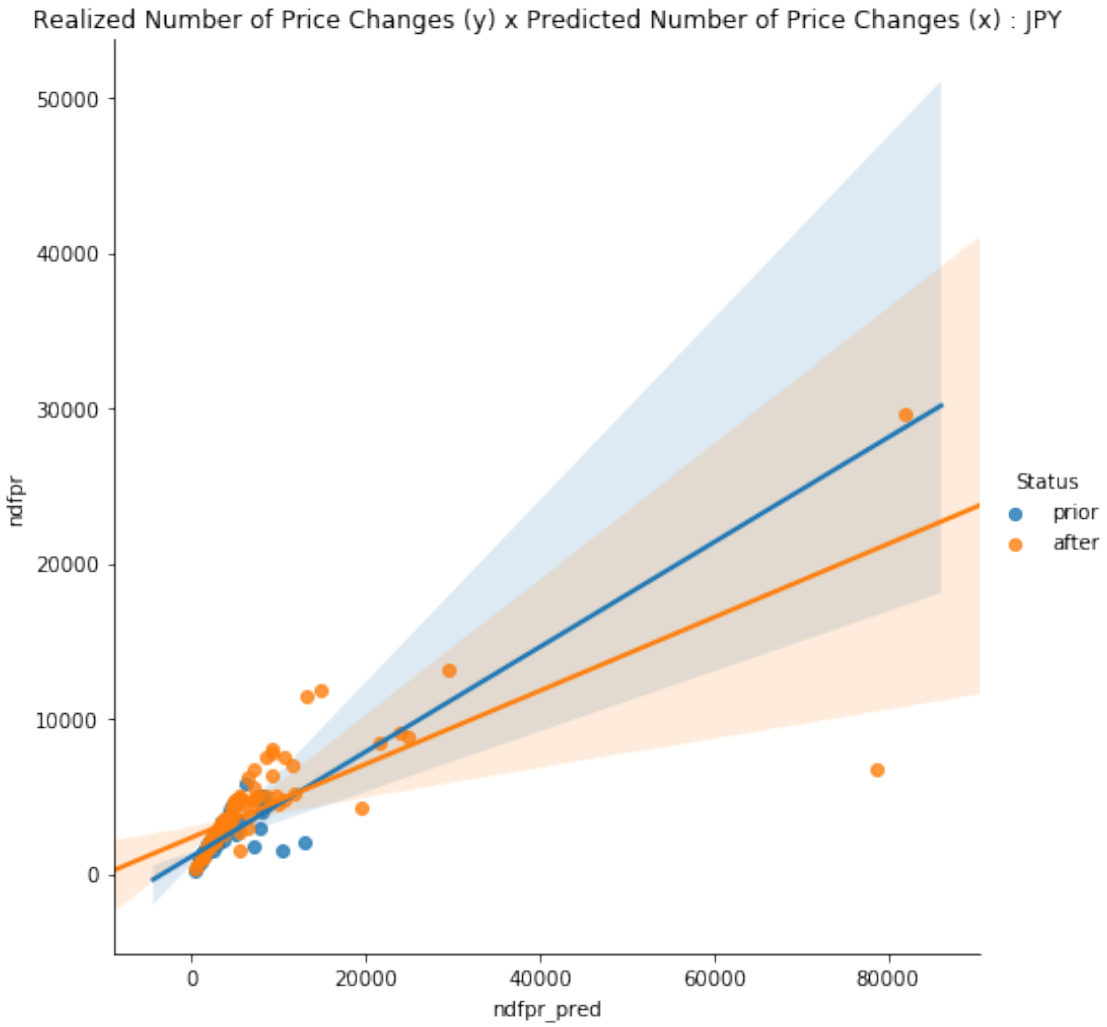
```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      M   R-squared:                       0.572
Model:                            OLS   Adj. R-squared:                  0.569
Method:                 Least Squares   F-statistic:                     164.6
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.93e-24
Time:                        16:11:24   Log-Likelihood:                -47.153
No. Observations:                 125   AIC:                             98.31
Df Residuals:                     123   BIC:                             104.0
Df Model:                           1
```

```
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         14.8451      0.408     36.360      0.000      14.037      15.653
rvxe           0.9043      0.070     12.831      0.000       0.765       1.044
==============================================================================
Omnibus:                       69.215   Durbin-Watson:                   1.222
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              304.490
Skew:                          -1.964   Prob(JB):                    7.60e-67
Kurtosis:                       9.560   Cond. No.                        76.5
==============================================================================

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

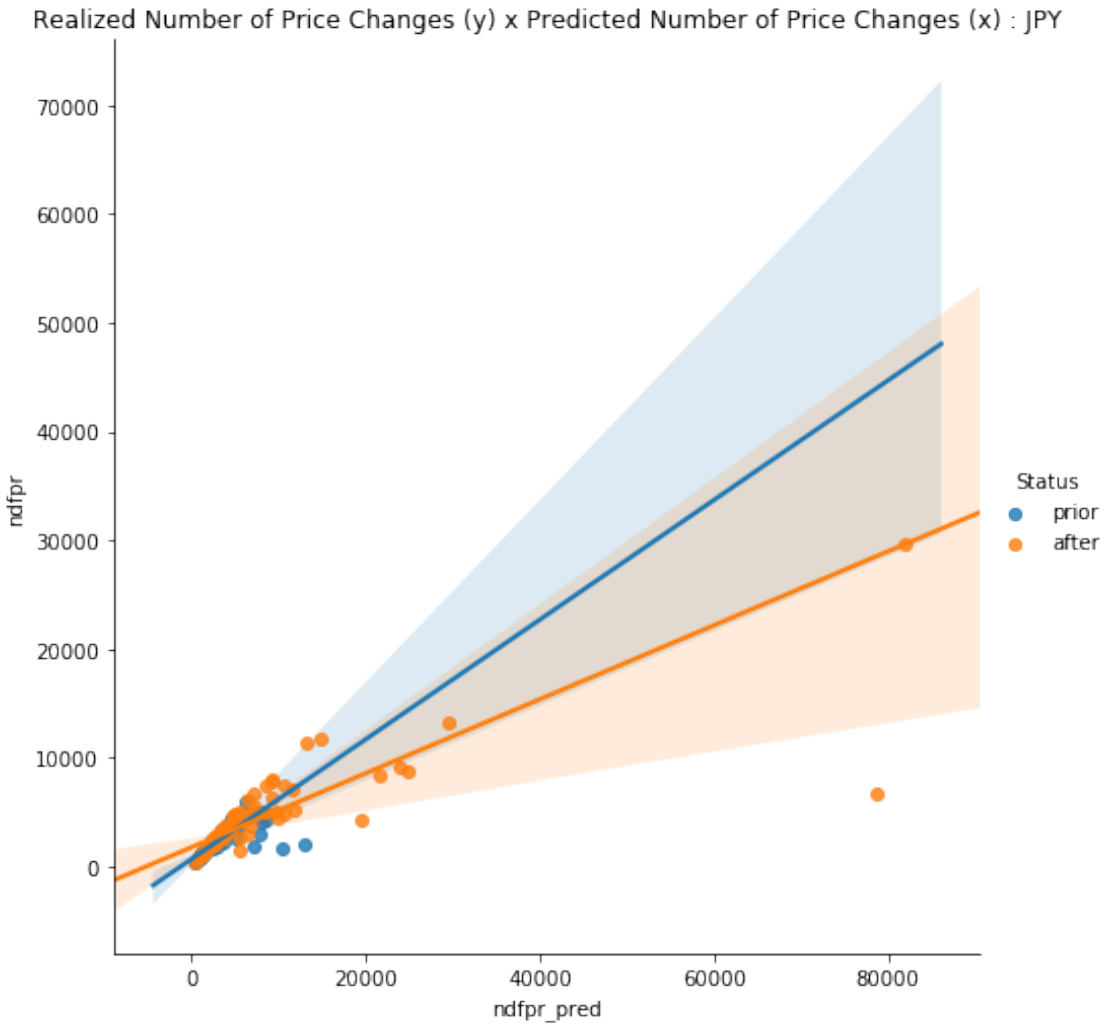
[115]:
```
cme.lin_reg_rob(AFTER_OB_UZ_STATS, 'rvxe', 'M', True)
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                      M   No. Observations:                  125
Model:                            RLM   Df Residuals:                      123
Method:                          IRLS   Df Model:                            1
Norm:                         HuberT
Scale Est.:                      mad
Cov Type:                         H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:24
No. Iterations:                    19
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         14.6011      0.311     46.884      0.000      13.991      15.211
rvxe           0.8556      0.054     15.915      0.000       0.750       0.961
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[116]:
```
cme.regr_plot(OB_UZ_STATS, 'ndfpr_pred', 'ndfpr',\
    'Realized Number of Price Changes (y) x Predicted Number of Price Changes␣
 ↪(x) : '+CURR)
```



[117]:
```
cme.regr_plot(OB_UZ_STATS, 'ndfpr_pred', 'ndfpr',\
    'Realized Number of Price Changes (y) x Predicted Number of Price Changes␣
↪(x) : '+CURR, True)
```



[118]:
```
cme.lin_reg(PRIOR_OB_UZ_STATS, 'ndfpr_pred', 'ndfpr')
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  ndfpr   R-squared:                       0.469
Model:                            OLS   Adj. R-squared:                  0.464
Method:                 Least Squares   F-statistic:                     99.66
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           3.30e-17
Time:                        16:11:38   Log-Likelihood:                -923.48
No. Observations:                 115   AIC:                             1851.
Df Residuals:                     113   BIC:                             1856.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

```
const       1137.3348    110.685     10.275      0.000     918.048    1356.622
ndfpr_pred     0.3372      0.034      9.983      0.000       0.270       0.404
==============================================================================
Omnibus:                       33.430   Durbin-Watson:                   1.205
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              219.570
Skew:                          -0.673   Prob(JB):                     2.09e-48
Kurtosis:                       9.634   Cond. No.                     5.19e+03
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 5.19e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

[119]:
```
cme.lin_reg_rob(PRIOR_OB_UZ_STATS, 'ndfpr_pred', 'ndfpr')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                  ndfpr   No. Observations:                  115
Model:                            RLM   Df Residuals:                      113
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:38
No. Iterations:                    24
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const        682.0525     47.193     14.452      0.000     589.556     774.549
ndfpr_pred     0.5505      0.014     38.218      0.000       0.522       0.579
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[120]:
```
cme.lin_reg(AFTER_OB_UZ_STATS, 'ndfpr_pred', 'ndfpr')
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                  ndfpr   R-squared:                       0.613
Model:                            OLS   Adj. R-squared:                  0.610
Method:                 Least Squares   F-statistic:                     194.7
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           4.10e-27
Time:                        16:11:38   Log-Likelihood:                -1126.7
No. Observations:                 125   AIC:                             2257.
```

```
Df Residuals:                      123   BIC:                             2263.
Df Model:                            1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       2355.3637    207.299     11.362      0.000    1945.028    2765.699
ndfpr_pred     0.2363      0.017     13.954      0.000       0.203       0.270
==============================================================================
Omnibus:                       90.852   Durbin-Watson:                   1.298
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             2561.497
Skew:                          -1.873   Prob(JB):                         0.00
Kurtosis:                      24.858   Cond. No.                     1.42e+04
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 1.42e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

[121]:
```
cme.lin_reg_rob(AFTER_OB_UZ_STATS, 'ndfpr_pred', 'ndfpr')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                   ndfpr   No. Observations:                  125
Model:                             RLM   Df Residuals:                      123
Method:                           IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:38
No. Iterations:                    15
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const       1757.2247     90.328     19.454      0.000    1580.185    1934.264
ndfpr_pred     0.3403      0.007     46.127      0.000       0.326       0.355
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[122]:
```
cme.regr_plot(OB_UZ_STATS, 'ndfpr', 'M',\
    'Number of Trades (y) x Number of Price Changes (x) : '+CURR)
```

[123]:
```
cme.regr_plot(OB_UZ_STATS, 'ndfpr', 'M',\
    'Number of Trades (y) x Number of Price Changes (x) : '+CURR, True)
```

[124]:
```
cme.lin_reg(PRIOR_OB_UZ_STATS, 'ndfpr', 'M')
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      M   R-squared:                       0.909
Model:                            OLS   Adj. R-squared:                  0.908
Method:                 Least Squares   F-statistic:                     1133.
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.00e-60
Time:                        16:11:44   Log-Likelihood:                -1032.0
No. Observations:                 115   AIC:                             2068.
Df Residuals:                     113   BIC:                             2073.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

```
const       1288.8412    394.509      3.267      0.001     507.247    2070.436
ndfpr          5.9288      0.176     33.656      0.000       5.580       6.278
==============================================================================
Omnibus:                        7.519   Durbin-Watson:                   1.181
Prob(Omnibus):                  0.023   Jarque-Bera (JB):               12.085
Skew:                          -0.220   Prob(JB):                      0.00238
Kurtosis:                       4.526   Cond. No.                     4.92e+03
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 4.92e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

[125]:
```
cme.lin_reg_rob(PRIOR_OB_UZ_STATS, 'ndfpr', 'M')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                      M   No. Observations:                  115
Model:                            RLM   Df Residuals:                      113
Method:                          IRLS   Df Model:                            1
Norm:                         HuberT
Scale Est.:                      mad
Cov Type:                         H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:44
No. Iterations:                     7
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const       1087.7832    353.087      3.081      0.002     395.745    1779.821
ndfpr          6.0329      0.158     38.264      0.000       5.724       6.342
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[126]:
```
cme.lin_reg(AFTER_OB_UZ_STATS, 'ndfpr', 'M')
```

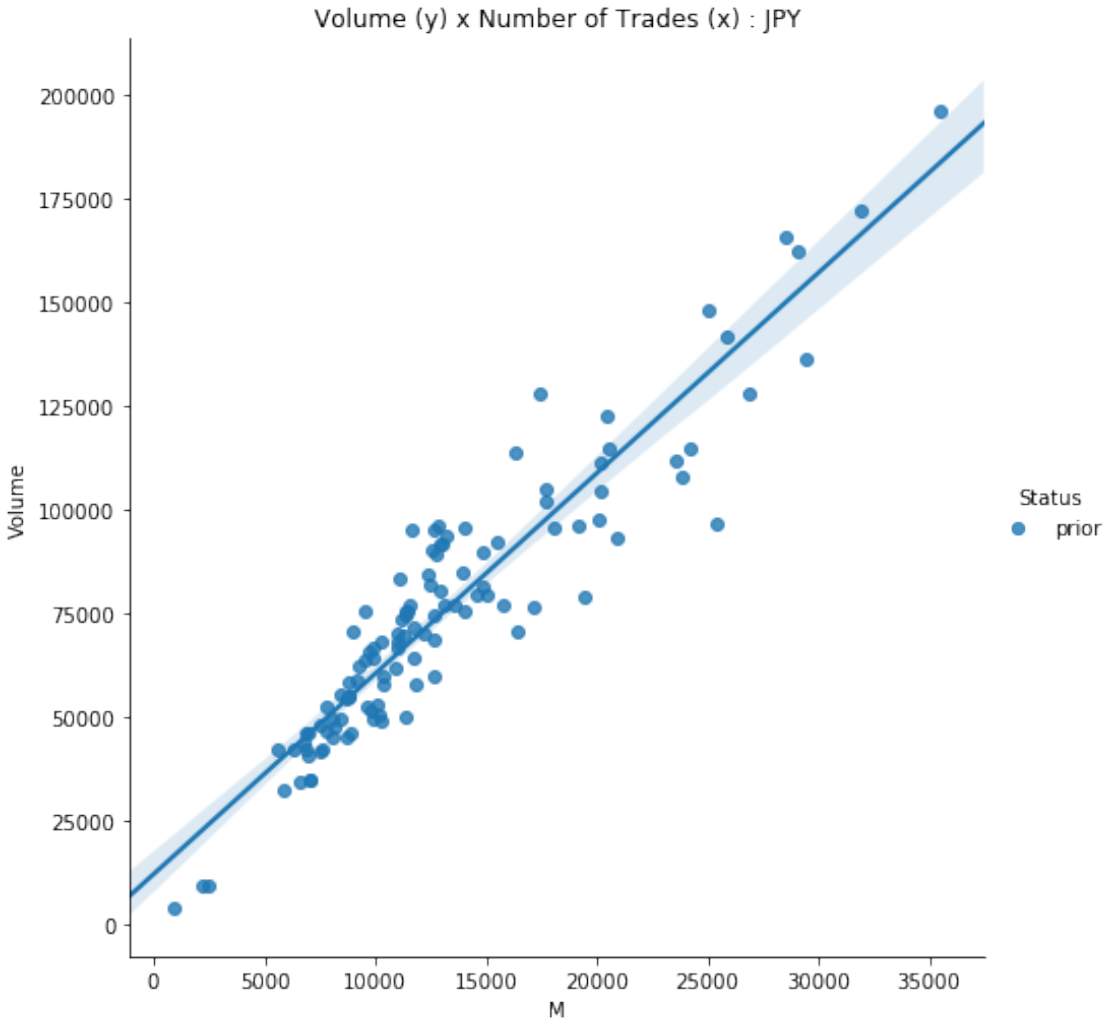
```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      M   R-squared:                       0.727
Model:                            OLS   Adj. R-squared:                  0.725
Method:                 Least Squares   F-statistic:                     328.1
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.65e-36
Time:                        16:11:44   Log-Likelihood:                -1236.4
No. Observations:                 125   AIC:                             2477.
```

```
Df Residuals:                      123   BIC:                             2482.
Df Model:                            1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       7900.9868    670.723     11.780      0.000    6573.333    9228.641
ndfpr          2.4437      0.135     18.113      0.000       2.177       2.711
==============================================================================
Omnibus:                       25.963   Durbin-Watson:                   1.047
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              204.287
Skew:                           0.153   Prob(JB):                     4.36e-45
Kurtosis:                       9.255   Cond. No.                     7.74e+03
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 7.74e+03. This might indicate that there are
strong multicollinearity or other numerical problems.
```

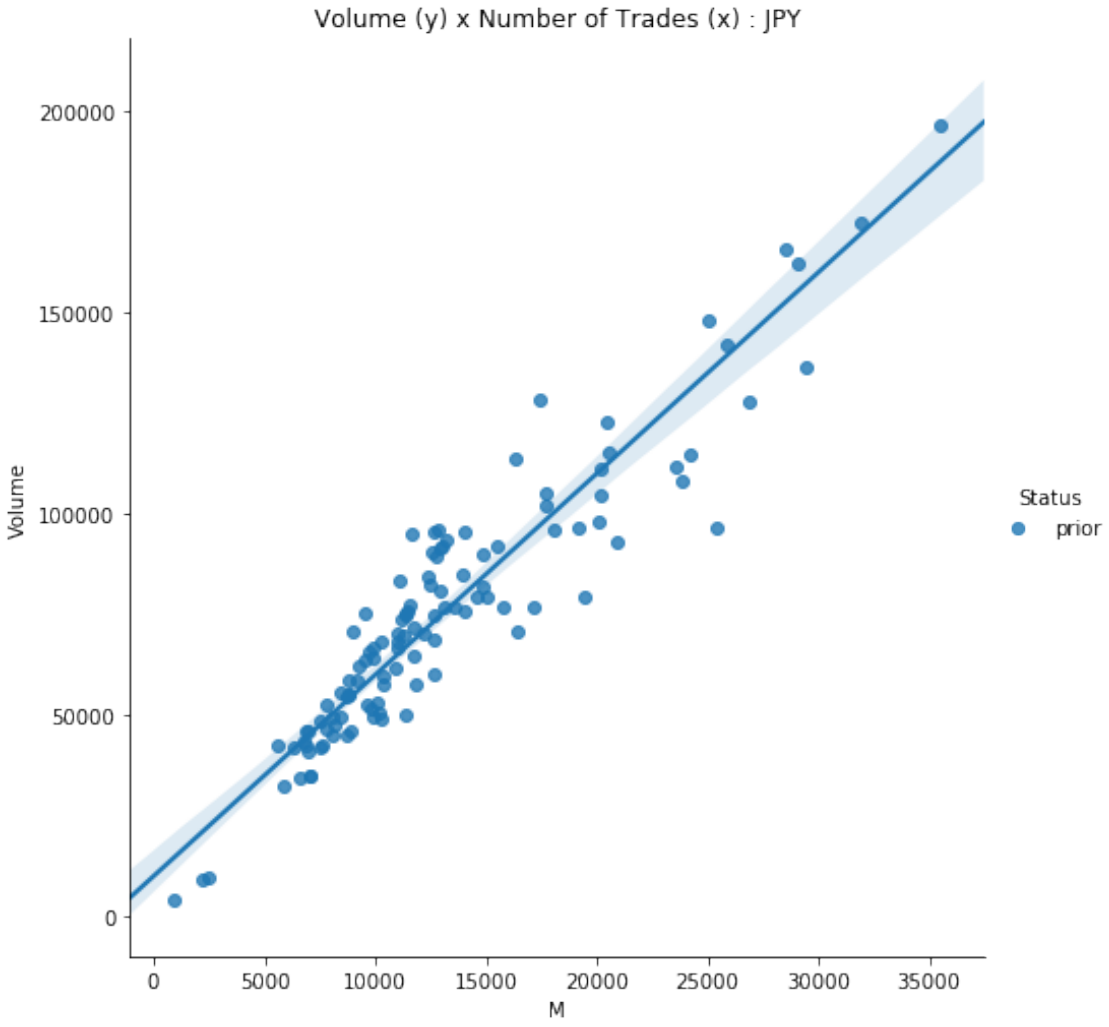
[127]:
```
cme.lin_reg_rob(AFTER_OB_UZ_STATS, 'ndfpr', 'M')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                      M   No. Observations:                  125
Model:                            RLM   Df Residuals:                      123
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:44
No. Iterations:                     9
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const       5883.7790    500.390     11.758      0.000    4903.033    6864.525
ndfpr          2.9554      0.101     29.362      0.000       2.758       3.153
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[128]:
```
cme.regr_plot(PRIOR_OB_UZ_STATS, 'M', 'Volume',\
    'Volume (y) x Number of Trades (x) : '+CURR)
```



[129]:
```
cme.regr_plot(PRIOR_OB_UZ_STATS, 'M', 'Volume',\
    'Volume (y) x Number of Trades (x) : '+CURR, True)
```



[130]:
```
cme.lin_reg(PRIOR_OB_UZ_STATS, 'M', 'Volume')
```

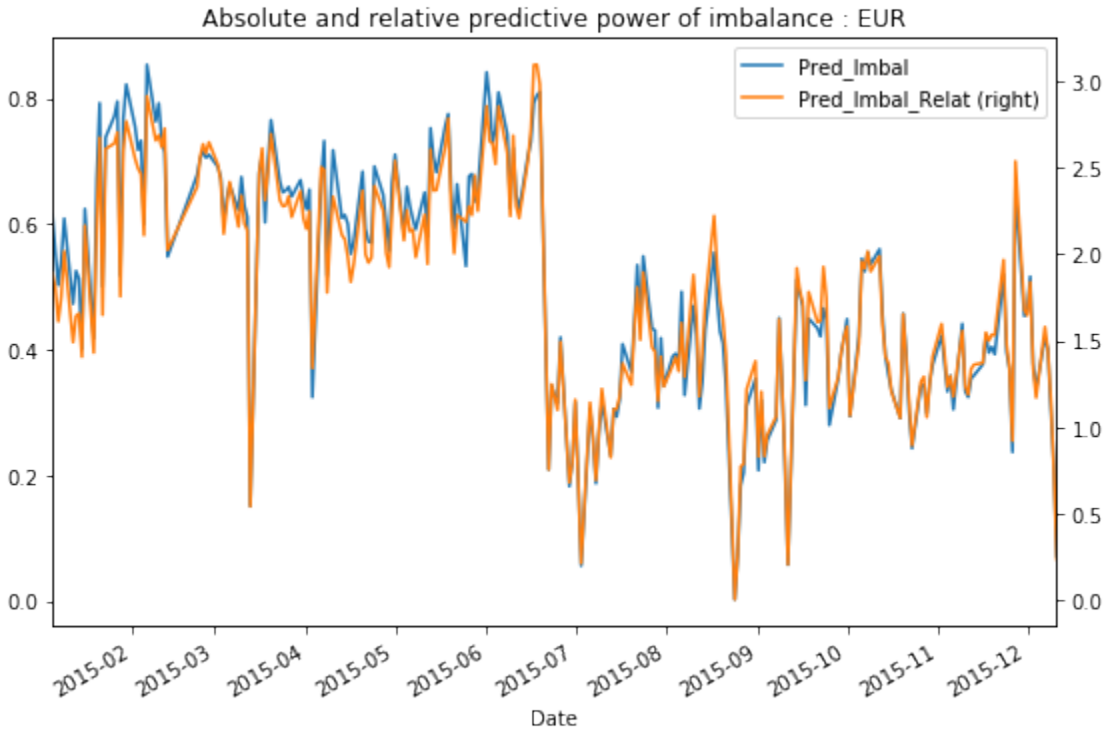
```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 Volume   R-squared:                       0.875
Model:                            OLS   Adj. R-squared:                  0.874
Method:                 Least Squares   F-statistic:                     789.6
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           8.24e-53
Time:                        16:11:47   Log-Likelihood:                -1239.5
No. Observations:                 115   AIC:                             2483.
Df Residuals:                     113   BIC:                             2488.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

```
const       1.209e+04   2507.128      4.822      0.000    7122.220    1.71e+04
M              4.8377      0.172     28.100      0.000       4.497       5.179
==============================================================================
Omnibus:                        1.366   Durbin-Watson:                   1.022
Prob(Omnibus):                  0.505   Jarque-Bera (JB):                0.897
Skew:                          -0.045   Prob(JB):                        0.638
Kurtosis:                       3.423   Cond. No.                     3.34e+04
==============================================================================

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly
specified.
[2] The condition number is large, 3.34e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

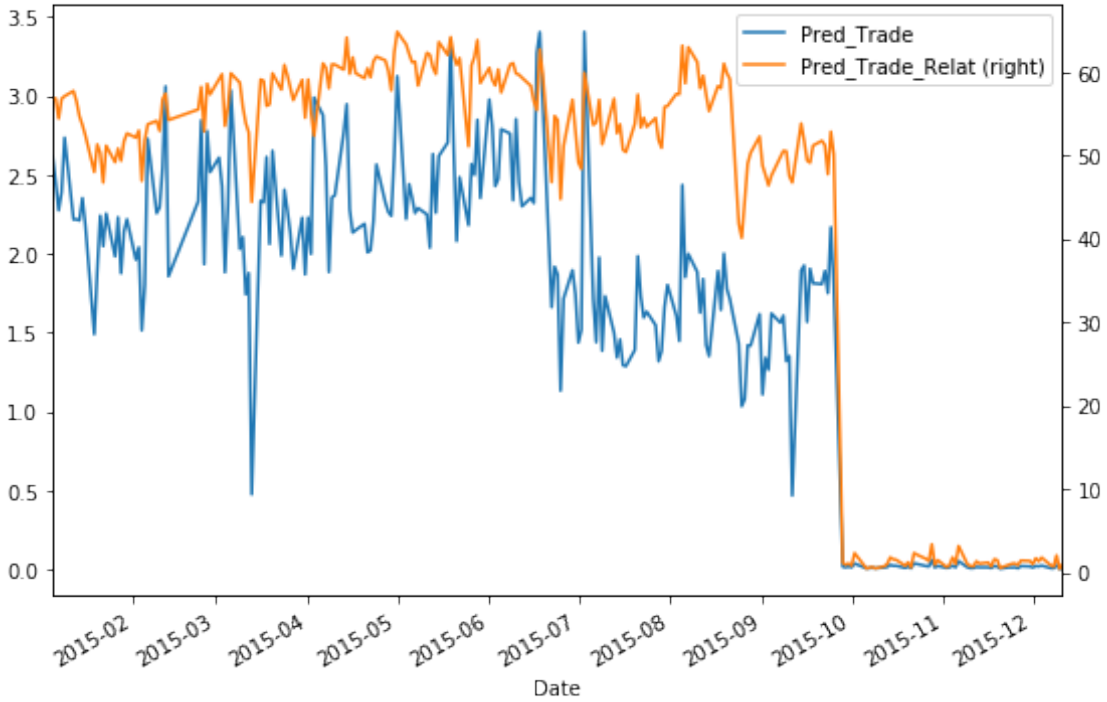
[131]:
```
cme.lin_reg_rob(PRIOR_OB_UZ_STATS, 'M', 'Volume')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                 Volume   No. Observations:                  115
Model:                            RLM   Df Residuals:                      113
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:47
No. Iterations:                     4
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const       9923.0892   2430.707      4.082      0.000    5158.992    1.47e+04
M              5.0035      0.167     29.977      0.000       4.676       5.331
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[132]:
```
IMBAL_STATS_TS.drop(columns=['eta1']).plot(secondary_y=['Pred_Imbal_Relat'],\
    figsize=(9,6), title='Absolute and relative predictive power of imbalance :␣
 ↪EUR');
```

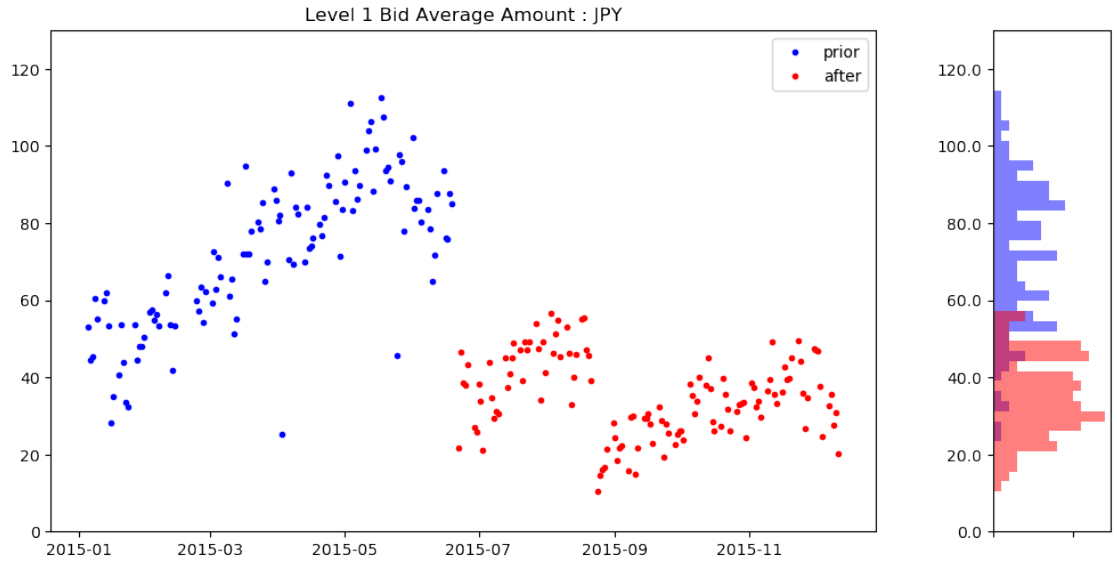


[133]:
```
TRADE_STATS_TS.plot(secondary_y=['Pred_Trade_Relat'], figsize=(9,6));
```

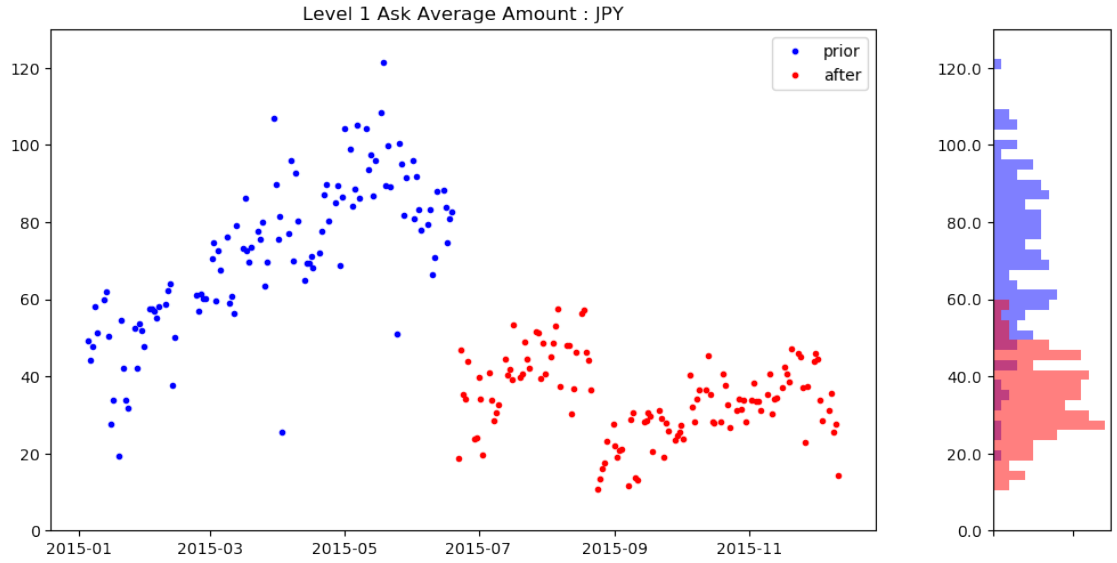


[134]:
```
OB_UZ_STATS_SPREADS = cme.spread_stats(OB_UZ_STATS)
```

[195]:
```
cme.time_series_hist_plot(OB_UZ_STATS_SPREADS, 'bid1qty',\
    'Level 1 Bid Average Amount : '+CURR, 0, 130, 50)
```



[197]:
```
cme.time_series_hist_plot(OB_UZ_STATS_SPREADS, 'ask1qty',\
    'Level 1 Ask Average Amount : '+CURR, 0, 130, 50)
```

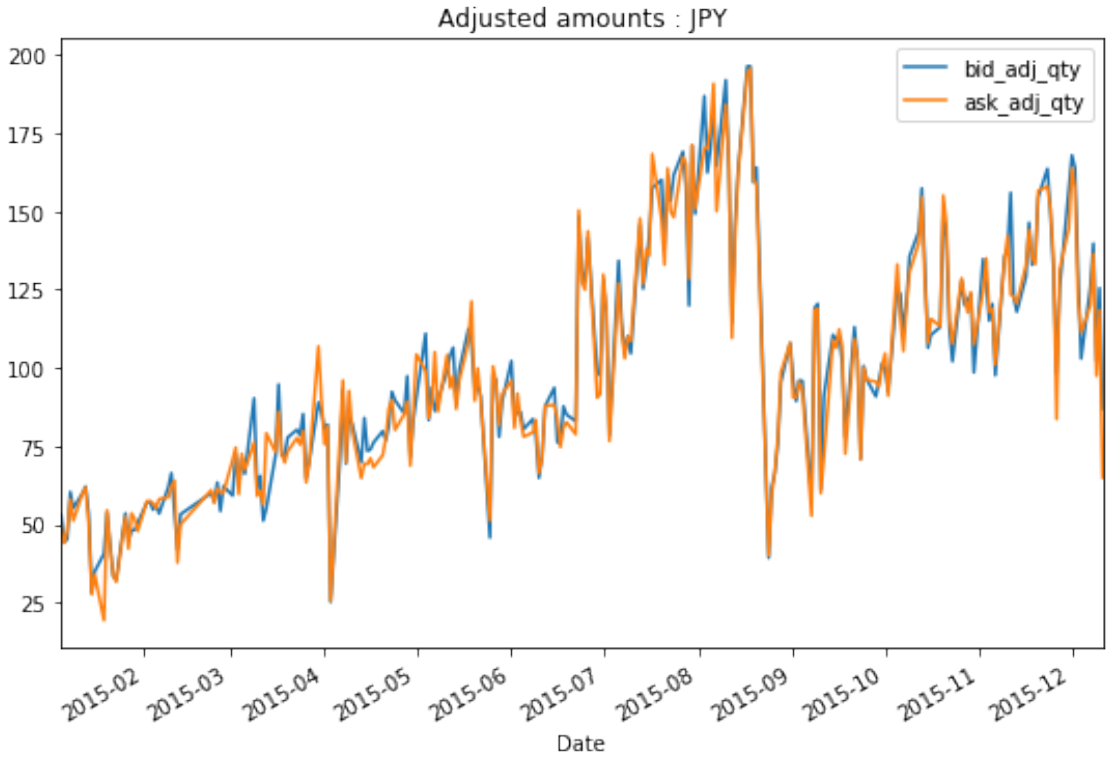


[137]:
```
OB_UZ_STATS_SPREADS[['bid1qty', 'ask1qty']].plot(figsize=(9,6));
```

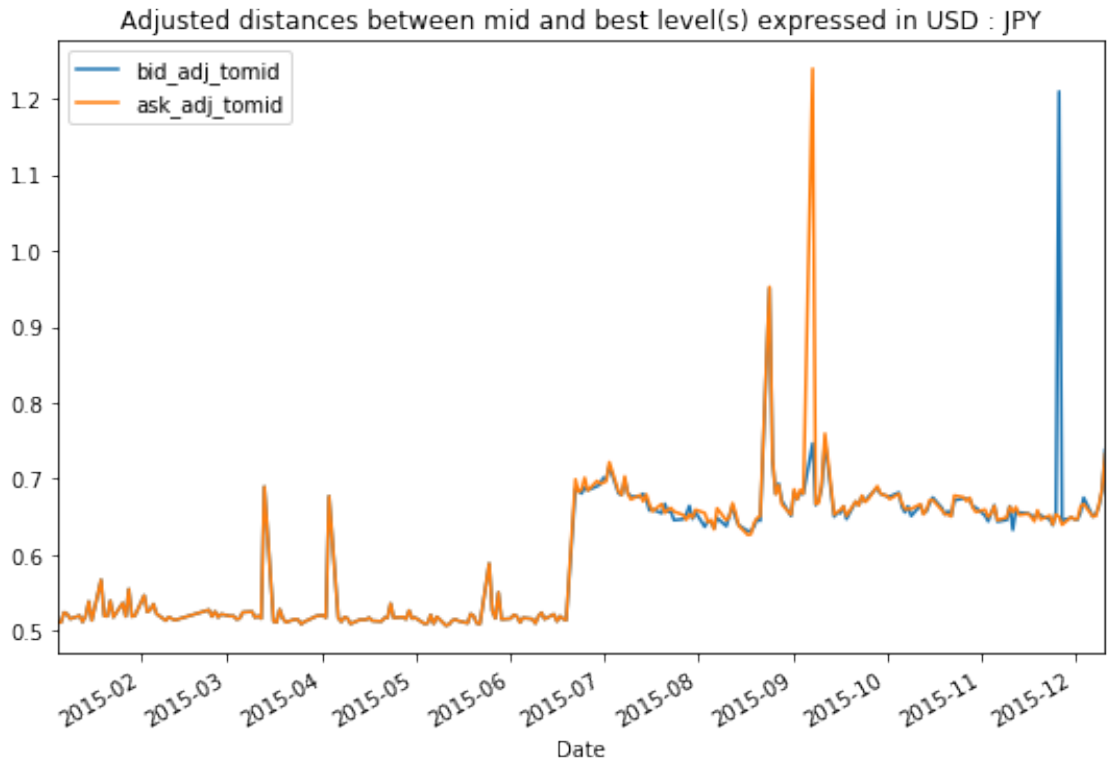
[138]:
```
OB_UZ_STATS_SPREADS[OB_UZ_STATS_SPREADS['Status'] == 'prior'][['bid1qty',␣
↪'ask1qty']].mean()/\
    OB_UZ_STATS_SPREADS[OB_UZ_STATS_SPREADS['Status'] == 'after'][['bid1qty',␣
↪'ask1qty']].mean()
```

[138]:
```
bid1qty    2.066217
ask1qty    2.104551
dtype: float64
```

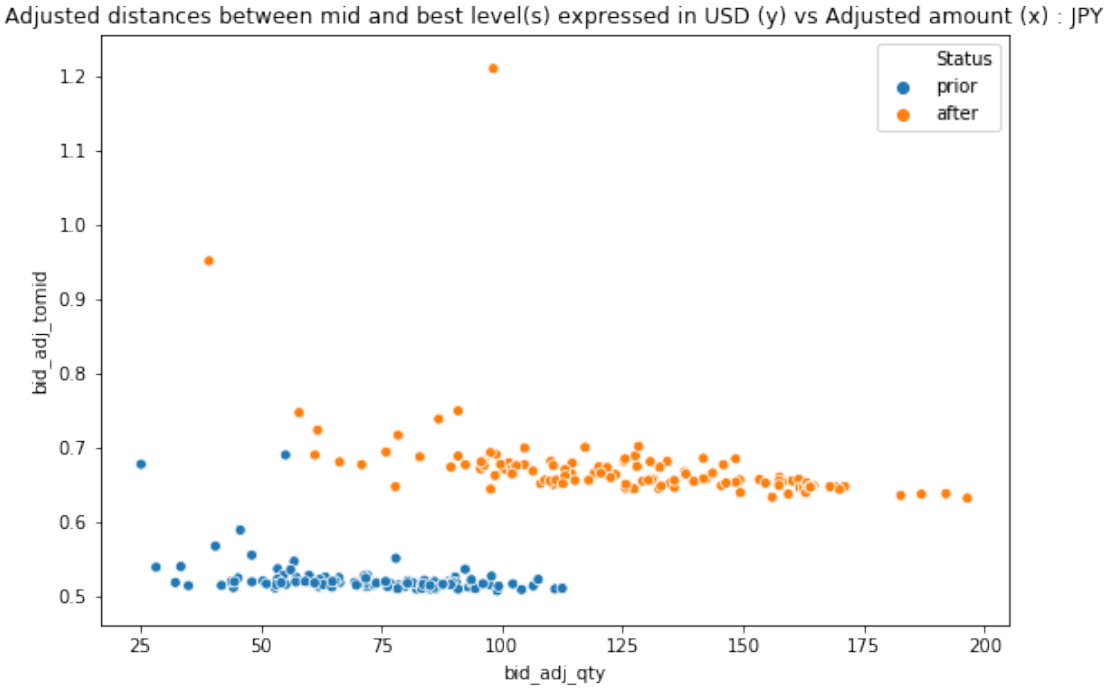
[139]:
```
OB_UZ_STATS_SPREADS[['bid_adj_qty', 'ask_adj_qty']].plot(figsize=(9,6),\
    title='Adjusted amounts : '+CURR);
```



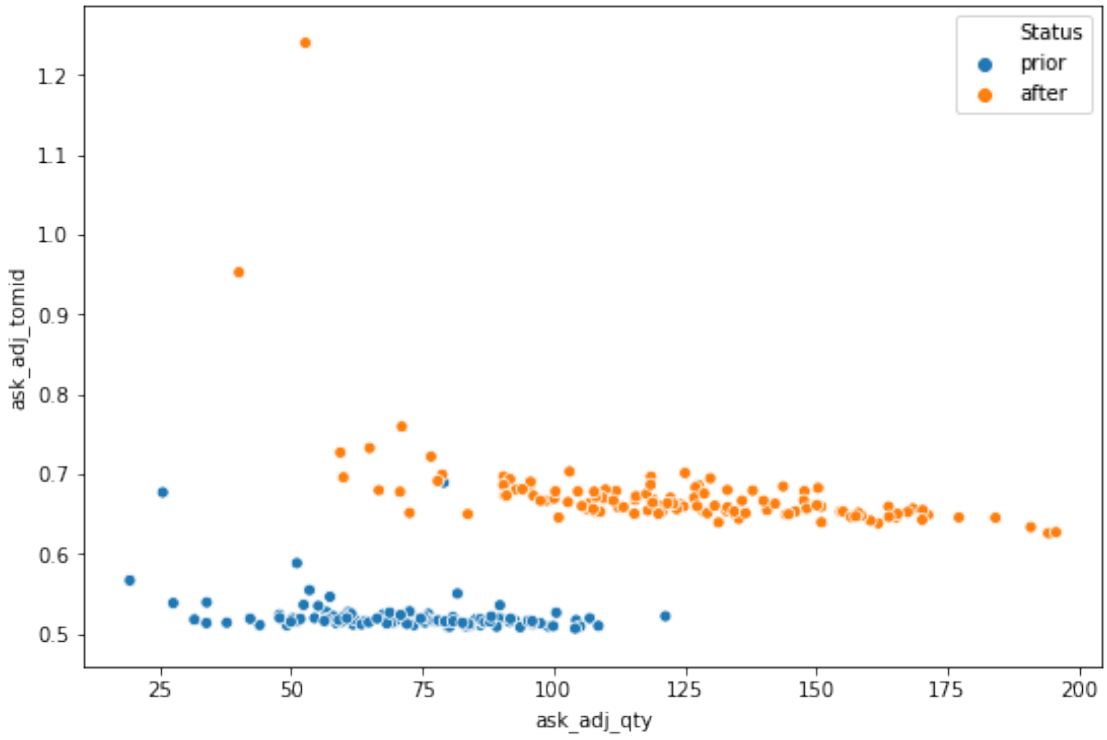
[140]:
```
OB_UZ_STATS_SPREADS[['bid_adj_tomid', 'ask_adj_tomid']].plot(figsize=(9,6),\
    title='Adjusted distances between mid and best level(s) expressed in USD :␣
↪'+CURR);
```



[141]:
```
plt.figure(figsize=(9, 6))
sns.scatterplot(x='bid_adj_qty', y='bid_adj_tomid',\
                hue='Status', data=OB_UZ_STATS_SPREADS);
plt.title('Adjusted distances between mid and best level(s) expressed in USD␣
 ↪(y) vs Adjusted amount (x) : '+CURR);
```



[142]:
```
plt.figure(figsize=(9, 6))
sns.scatterplot(x='ask_adj_qty', y='ask_adj_tomid',\
                hue='Status', data=OB_UZ_STATS_SPREADS);
```



2.8.1 Costs

[143]:
```
PRIOR_MEAN_COST = cme.cost_mean(PRIOR_COST_STATS, 100)
```

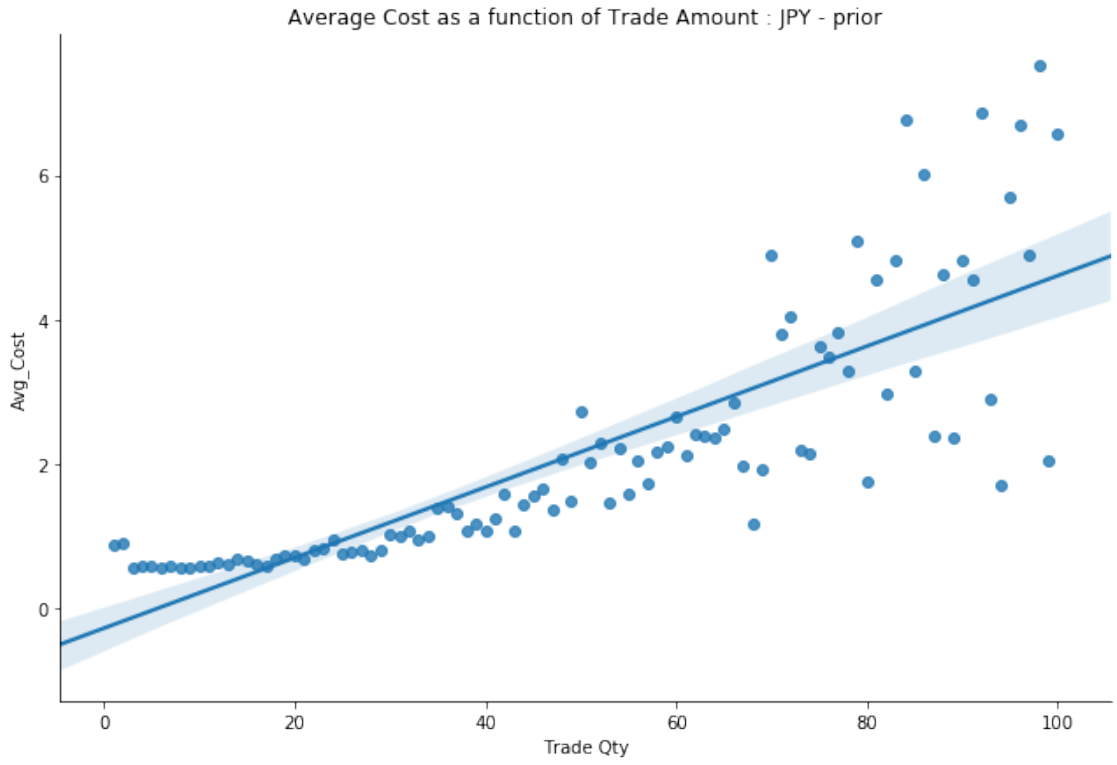
[144]:
```
PRIOR_MEAN_COST['Status'] = 'prior'
```

[145]:
```
AFTER_MEAN_COST = cme.cost_mean(AFTER_COST_STATS, 100)
```

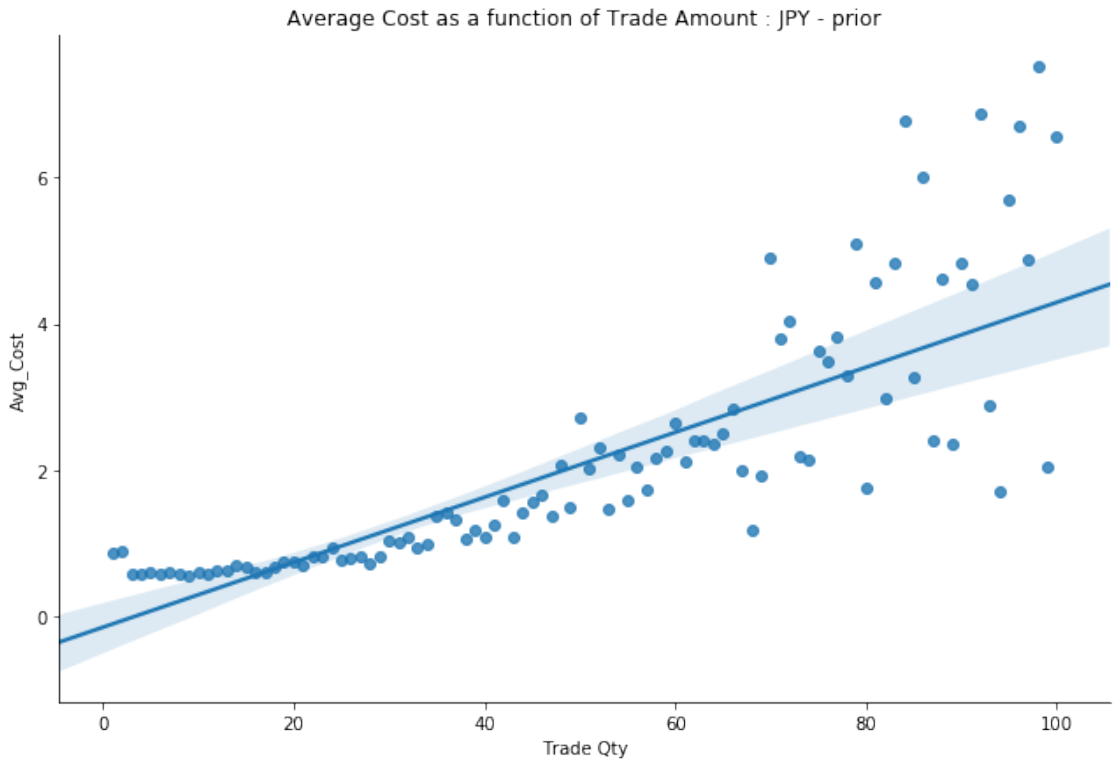
[146]:
```
AFTER_MEAN_COST['Status'] = 'after'
```

[147]:
```
MEAN_COST_STATS = pd.concat([PRIOR_MEAN_COST, AFTER_MEAN_COST], sort=False)
```

[148]:
```
sns.lmplot(x='Trade Qty', y='Avg_Cost', data=PRIOR_MEAN_COST.reset_index(),\
           height=6, aspect=1.5);
plt.title('Average Cost as a function of Trade Amount : '+CURR+' - prior');
```



[149]:
```
sns.lmplot(x='Trade Qty', y='Avg_Cost', data=PRIOR_MEAN_COST.reset_index(),\
           height=6, aspect=1.5, robust=True);
plt.title('Average Cost as a function of Trade Amount : '+CURR+' - prior');
```



[150]:
```
cme.lin_reg(cme.cost_mean(PRIOR_COST_STATS, 50).reset_index(), 'Trade Qty',␣
↪'Avg_Cost')
```

OLS Regression Results

Dep. Variable:	Avg_Cost	R-squared:	0.669
Model:	OLS	Adj. R-squared:	0.662
Method:	Least Squares	F-statistic:	96.89
Date:	Wed, 09 Oct 2019	Prob (F-statistic):	4.24e-13
Time:	16:11:58	Log-Likelihood:	-1.4328
No. Observations:	50	AIC:	6.866
Df Residuals:	48	BIC:	10.69
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.3491	0.073	4.784	0.000	0.202	0.496
Trade Qty	0.0245	0.002	9.843	0.000	0.020	0.030

Omnibus:	43.627	Durbin-Watson:	1.118
Prob(Omnibus):	0.000	Jarque-Bera (JB):	161.855
Skew:	2.326	Prob(JB):	7.14e-36

```
Kurtosis:                       10.487   Cond. No.                         59.5
==============================================================================

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

[151]:
```
cme.lin_reg_rob(cme.cost_mean(PRIOR_COST_STATS, 50).reset_index(), 'Trade Qty',␣
↪'Avg_Cost')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:               Avg_Cost   No. Observations:                   50
Model:                            RLM   Df Residuals:                       48
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:58
No. Iterations:                    24
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.3657      0.051      7.125      0.000       0.265       0.466
Trade Qty      0.0223      0.002     12.750      0.000       0.019       0.026
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[152]:
```
cme.lin_reg(cme.cost_mean(PRIOR_COST_STATS, 100).reset_index(), 'Trade Qty',␣
↪'Avg_Cost')
```

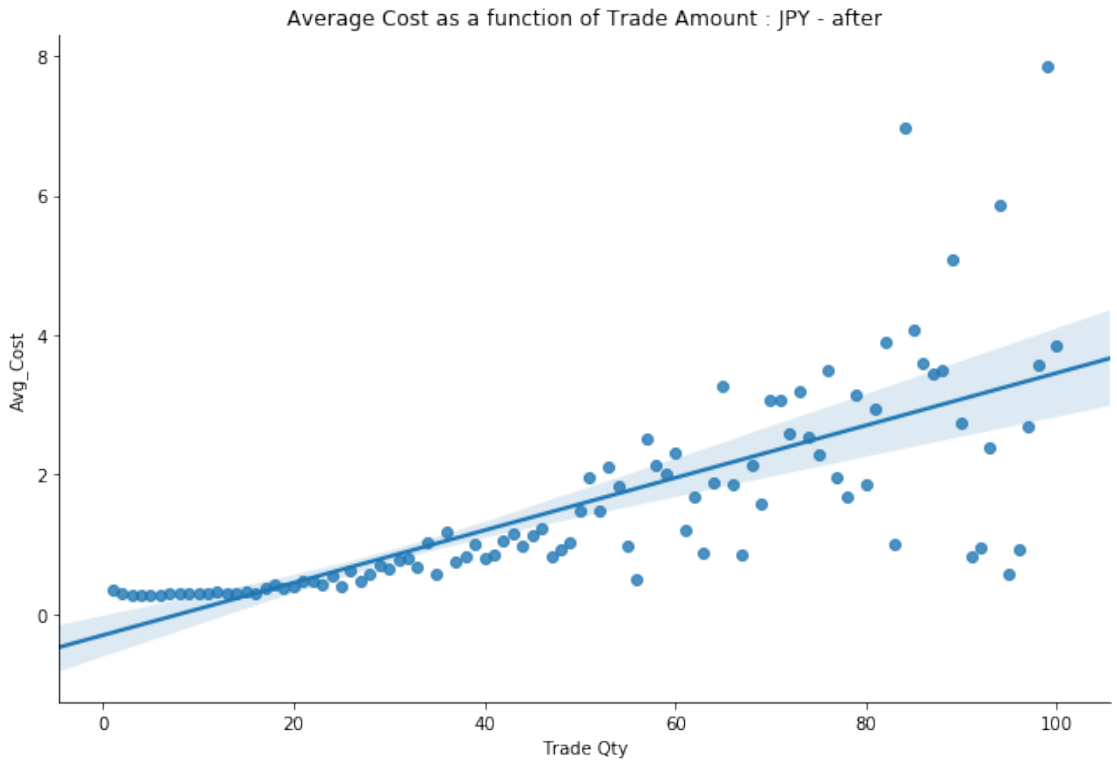
```
                            OLS Regression Results
==============================================================================
Dep. Variable:               Avg_Cost   R-squared:                       0.675
Model:                            OLS   Adj. R-squared:                  0.671
Method:                 Least Squares   F-statistic:                     203.1
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.26e-25
Time:                        16:11:58   Log-Likelihood:                -139.84
No. Observations:                 100   AIC:                             283.7
Df Residuals:                      98   BIC:                             288.9
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

```
const         -0.2767      0.199     -1.387      0.168      -0.672       0.119
Trade Qty      0.0489      0.003     14.252      0.000       0.042       0.056
==============================================================================
Omnibus:                       11.623   Durbin-Watson:                   2.148
Prob(Omnibus):                  0.003   Jarque-Bera (JB):               16.776
Skew:                           0.535   Prob(JB):                     0.000228
Kurtosis:                       4.698   Cond. No.                         117.
==============================================================================

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

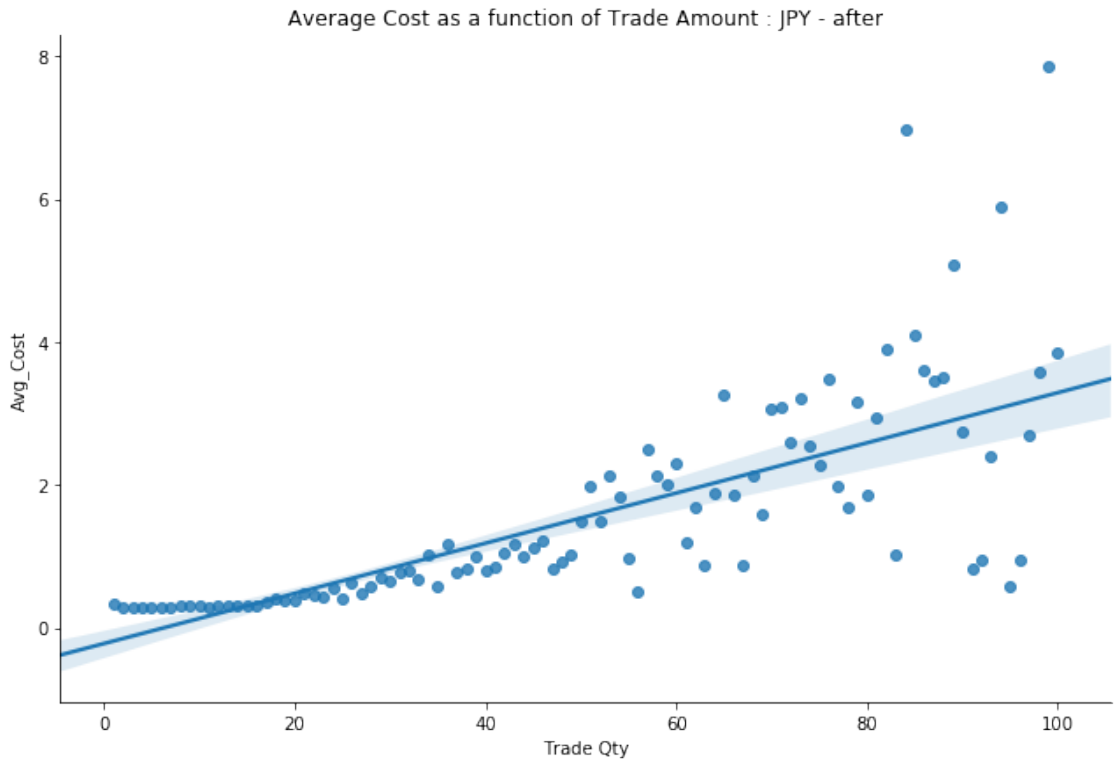
[153]:
```
cme.lin_reg_rob(cme.cost_mean(PRIOR_COST_STATS, 100).reset_index(), 'Trade␣
↪Qty', 'Avg_Cost')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                Avg_Cost   No. Observations:                  100
Model:                            RLM   Df Residuals:                       98
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:11:58
No. Iterations:                    37
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.1532      0.146     -1.050      0.294      -0.439       0.133
Trade Qty      0.0445      0.003     17.721      0.000       0.040       0.049
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[154]:
```
sns.lmplot(x='Trade Qty', y='Avg_Cost', data=AFTER_MEAN_COST.reset_index(),\
           height=6, aspect=1.5);
plt.title('Average Cost as a function of Trade Amount : '+CURR+' - after');
```



[155]:
```
sns.lmplot(x='Trade Qty', y='Avg_Cost', data=AFTER_MEAN_COST.reset_index(),\
           height=6, aspect=1.5, robust=True);
plt.title('Average Cost as a function of Trade Amount : '+CURR+' - after');
```



[156]:
```
cme.lin_reg(cme.cost_mean(AFTER_COST_STATS, 50).reset_index(), 'Trade Qty',␣
↪'Avg_Cost')
```

OLS Regression Results

Dep. Variable:	Avg_Cost	R-squared:	0.831
Model:	OLS	Adj. R-squared:	0.828
Method:	Least Squares	F-statistic:	236.4
Date:	Wed, 09 Oct 2019	Prob (F-statistic):	3.58e-20
Time:	16:12:05	Log-Likelihood:	30.399
No. Observations:	50	AIC:	-56.80
Df Residuals:	48	BIC:	-52.97
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0992	0.039	2.570	0.013	0.022	0.177
Trade Qty	0.0203	0.001	15.374	0.000	0.018	0.023

Omnibus:	6.257	Durbin-Watson:	1.786
Prob(Omnibus):	0.044	Jarque-Bera (JB):	5.337
Skew:	0.775	Prob(JB):	0.0694

```
Kurtosis:                       3.398   Cond. No.                         59.5
==============================================================================

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

[157]:
```
cme.lin_reg_rob(cme.cost_mean(AFTER_COST_STATS, 50).reset_index(), 'Trade Qty',␣
↪'Avg_Cost')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:               Avg_Cost   No. Observations:                   50
Model:                            RLM   Df Residuals:                       48
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:12:05
No. Iterations:                    21
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.1086      0.036      2.976      0.003       0.037       0.180
Trade Qty      0.0194      0.001     15.583      0.000       0.017       0.022
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

[158]:
```
cme.lin_reg(cme.cost_mean(AFTER_COST_STATS, 100).reset_index(), 'Trade Qty',␣
↪'Avg_Cost')
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:               Avg_Cost   R-squared:                       0.548
Model:                            OLS   Adj. R-squared:                  0.543
Method:                 Least Squares   F-statistic:                     118.6
Date:                Wed, 09 Oct 2019   Prob (F-statistic):           1.41e-18
Time:                        16:12:05   Log-Likelihood:                -140.55
No. Observations:                 100   AIC:                             285.1
Df Residuals:                      98   BIC:                             290.3
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
```

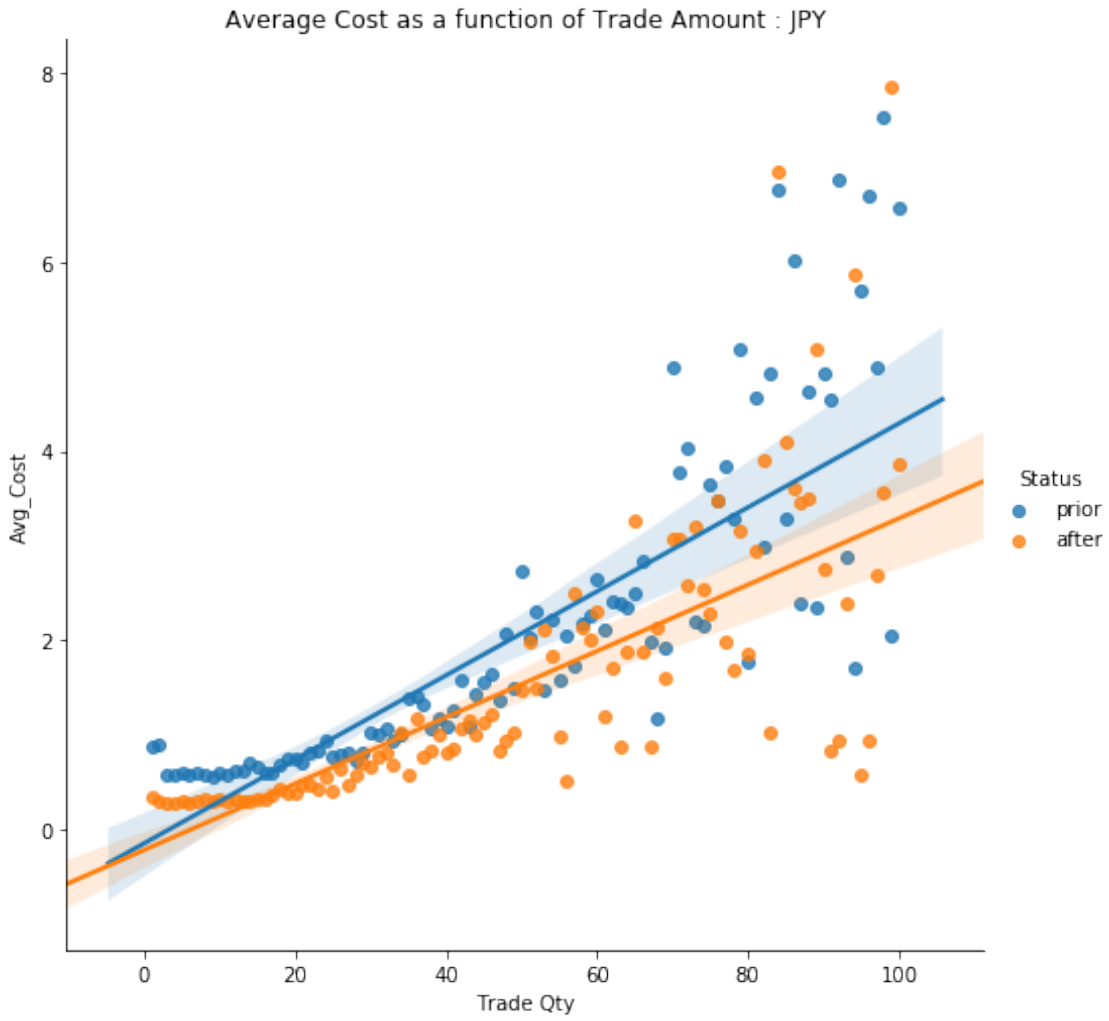
```
const         -0.3018      0.201     -1.503      0.136      -0.700       0.097
Trade Qty      0.0376      0.003     10.893      0.000       0.031       0.044
==============================================================================
Omnibus:                       42.575   Durbin-Watson:                   1.804
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              207.085
Skew:                           1.245   Prob(JB):                     1.08e-45
Kurtosis:                       9.595   Cond. No.                         117.
==============================================================================

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

[159]:
```
cme.lin_reg_rob(cme.cost_mean(AFTER_COST_STATS, 100).reset_index(), 'Trade␣
↪Qty', 'Avg_Cost')
```

```
                    Robust linear Model Regression Results
==============================================================================
Dep. Variable:                Avg_Cost   No. Observations:                  100
Model:                            RLM   Df Residuals:                       98
Method:                          IRLS   Df Model:                            1
Norm:                          HuberT
Scale Est.:                       mad
Cov Type:                          H1
Date:                Wed, 09 Oct 2019
Time:                        16:12:05
No. Iterations:                    24
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const         -0.2237      0.117     -1.917      0.055      -0.452       0.005
Trade Qty      0.0351      0.002     17.504      0.000       0.031       0.039
==============================================================================

If the model instance has been used for another fit with different fit
parameters, then the fit options might not be the correct ones anymore .
```

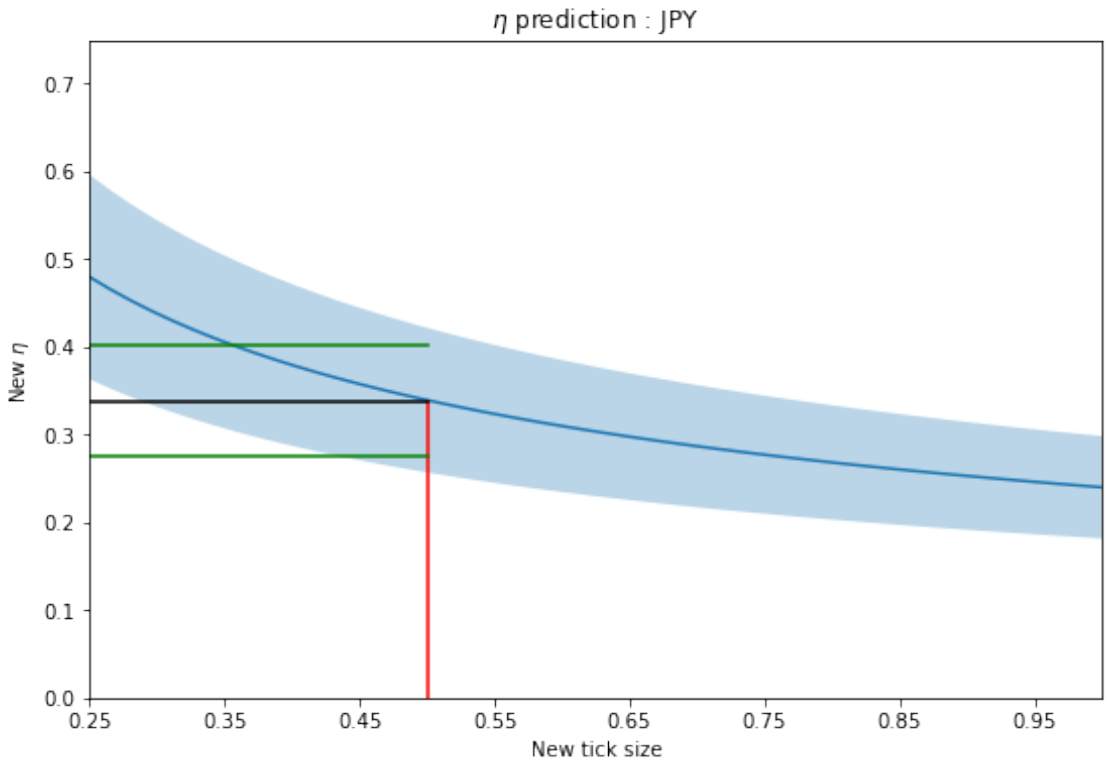
[160]:
```
cme.regr_plot(MEAN_COST_STATS.reset_index(), 'Trade Qty', 'Avg_Cost',\
    'Average Cost as a function of Trade Amount : '+CURR)
```

[161]:
```
cme.regr_plot(MEAN_COST_STATS.reset_index(), 'Trade Qty', 'Avg_Cost',\
    'Average Cost as a function of Trade Amount : '+CURR, True)
```



2.9 Eta prediction

[162]:
```
cme.plot_eta(TICK_PRIOR, TICK_AFTER,\
    TABLE_MATHIEU.loc['prior']['eta1'], TABLE_MATHIEU.loc['after']['eta1'],\
    TABLE_MATHIEU_ERR.loc['prior']['eta1'], TABLE_MATHIEU_ERR.
↪loc['after']['eta1'],\
    CURR)
```


η prediction : JPY
New η
New tick size
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.25
0.35
0.45
0.55
0.65
0.75
0.85
0.95

[ ]: