EconQuizFinal

November 11, 2025

[1]:
```
import pandas as pd  # <--- You need this import
import statsmodels.formula.api as smf
import matplotlib.pyplot as plt
```

[2]:
```
# This is the missing step. You must load your CSV into a DataFrame.
# We'll use the mock data example:
data = pd.read_csv("/Users/utkarshtyagi/Downloads/Q2_data_did.csv")
```

[4]:
```
model=smf.ols('hours_worked ~ treated * post + age + female +␣
 ↪sector',data=data).fit(cov_type='HC1')
```

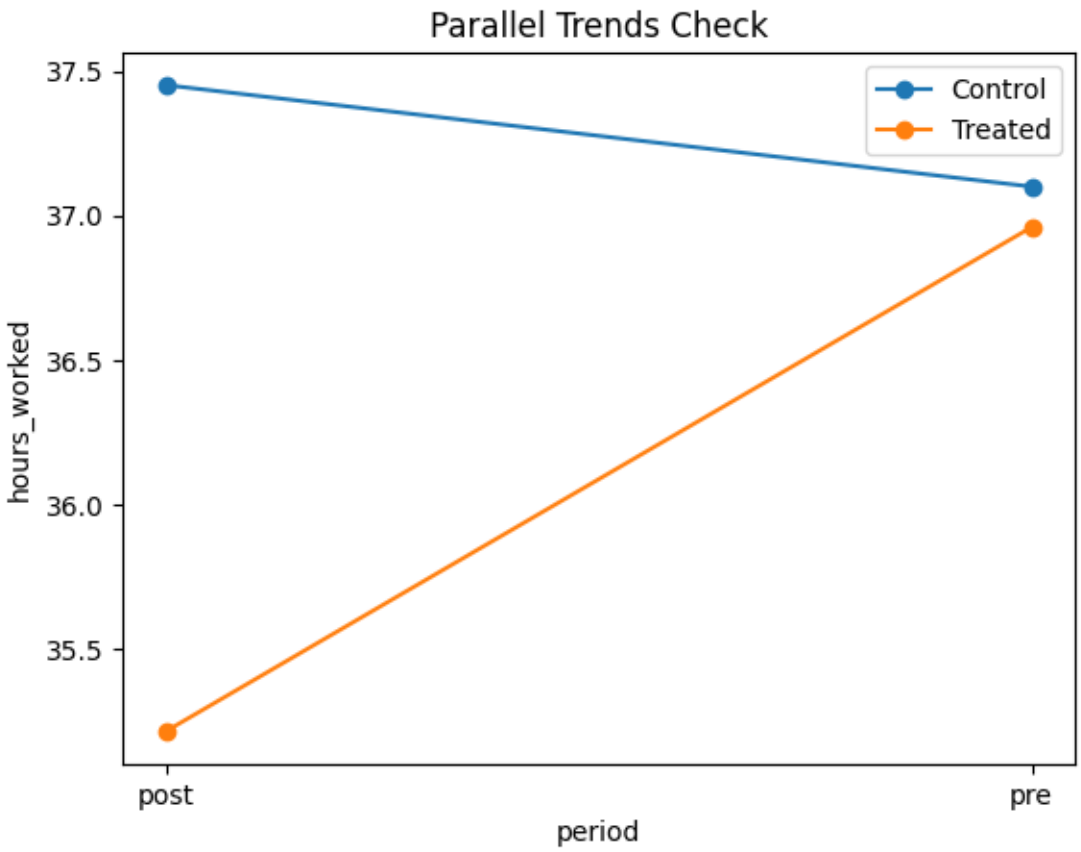
[5]:
```
print(model.summary())
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:           hours_worked   R-squared:                       0.313
Model:                            OLS   Adj. R-squared:                  0.306
Method:                 Least Squares   F-statistic:                     50.55
Date:                Tue, 11 Nov 2025   Prob (F-statistic):           3.31e-50
Time:                        18:31:24   Log-Likelihood:                -1275.1
No. Observations:                 600   AIC:                             2564.
Df Residuals:                     593   BIC:                             2595.
Df Model:                           6
Covariance Type:                  HC1
==============================================================================
                 coef    std err          z      P>|z|      [0.025
0.975]
------------------------------------------------------------------------------
Intercept     40.4576      0.323    125.450      0.000      39.826
41.090
sector[T.Retail]    -0.4509      0.166     -2.724      0.006      -0.775
-0.126
treated       -0.0574      0.244     -0.236      0.814      -0.535
0.420
post           0.3516      0.231      1.520      0.128      -0.102
0.805
```

```
treated:post        -2.0971      0.333     -6.300      0.000      -2.749
-1.445
age                 -0.0656      0.007     -9.493      0.000      -0.079
-0.052
female              -1.2183      0.169     -7.219      0.000      -1.549
-0.888
==============================================================================
Omnibus:                        0.617   Durbin-Watson:                   2.054
Prob(Omnibus):                  0.735   Jarque-Bera (JB):                0.713
Skew:                           0.027   Prob(JB):                        0.700
Kurtosis:                       2.840   Cond. No.                         220.
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)
```

[7]:
```
# Parallel trends plot
avg_trends = data.groupby(['period', 'treated'])['hours_worked'].mean().
 ↪reset_index()
plt.plot(avg_trends[avg_trends['treated']==0]['period'],␣
 ↪avg_trends[avg_trends['treated']==0]['hours_worked'], marker='o',␣
 ↪label='Control')
plt.plot(avg_trends[avg_trends['treated']==1]['period'],␣
 ↪avg_trends[avg_trends['treated']==1]['hours_worked'], marker='o',␣
 ↪label='Treated')
plt.legend()
plt.xlabel("period")
plt.ylabel("hours_worked")
plt.title("Parallel Trends Check")
plt.show()
```



[11]:
```
# --- 5. ROBUSTNESS CHECK (Add controls) ---
controls = [' age ', ' female ', ' sector']    # replace with your controls
robust_formula = f" hours_worked ~ treated* post + {' + '.join(controls)}"
did_robust = smf.ols(robust_formula, data=data).fit(cov_type='HC1')
print("\n================ ROBUSTNESS (DiD + CONTROLS) ================")
print(did_robust.summary())
```

```
================ ROBUSTNESS (DiD + CONTROLS) ================
                            OLS Regression Results
==============================================================================
Dep. Variable:           hours_worked   R-squared:                       0.313
Model:                            OLS   Adj. R-squared:                  0.306
Method:                 Least Squares   F-statistic:                     50.55
Date:                Tue, 11 Nov 2025   Prob (F-statistic):           3.31e-50
Time:                        18:37:20   Log-Likelihood:                -1275.1
No. Observations:                 600   AIC:                             2564.
Df Residuals:                     593   BIC:                             2595.
Df Model:                           6
Covariance Type:                  HC1
```

```
==============================================================================
====
                       coef    std err          z      P>|z|      [0.025
0.975]
------------------------------------------------------------------------------
----
Intercept           40.4576      0.323    125.450      0.000      39.826
41.090
sector[T.Retail]    -0.4509      0.166     -2.724      0.006      -0.775
-0.126
treated             -0.0574      0.244     -0.236      0.814      -0.535
0.420
post                 0.3516      0.231      1.520      0.128      -0.102
0.805
treated:post        -2.0971      0.333     -6.300      0.000      -2.749
-1.445
age                 -0.0656      0.007     -9.493      0.000      -0.079
-0.052
female              -1.2183      0.169     -7.219      0.000      -1.549
-0.888
==============================================================================
Omnibus:                        0.617   Durbin-Watson:                   2.054
Prob(Omnibus):                  0.735   Jarque-Bera (JB):                0.713
Skew:                           0.027   Prob(JB):                        0.700
Kurtosis:                       2.840   Cond. No.                         220.
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)
```