In [1]:
```python
from __future__ import division
import pandas as pd
import numpy as np    # numerical operations
import statsmodels.api as sm
import statsmodels.formula.api as smf
import os
import matplotlib.pyplot as plt
```

```
In [2]:  !pip install linearmodels
         import linearmodels as lm

         from linearmodels import PanelOLS
         from linearmodels import RandomEffects
         from linearmodels import FirstDifferenceOLS
```

Requirement already satisfied: linearmodels in ./anaconda3/lib/python3.
11/site-packages (5.3)
Requirement already satisfied: numpy>=1.19.0 in ./anaconda3/lib/python
3.11/site-packages (from linearmodels) (1.24.3)
Requirement already satisfied: pandas>=1.1.0 in ./anaconda3/lib/python
3.11/site-packages (from linearmodels) (1.5.3)
Requirement already satisfied: scipy>=1.5.0 in ./anaconda3/lib/python3.
11/site-packages (from linearmodels) (1.10.1)
Requirement already satisfied: statsmodels>=0.12.0 in ./anaconda3/lib/p
ython3.11/site-packages (from linearmodels) (0.14.0)
Requirement already satisfied: mypy-extensions>=0.4 in ./anaconda3/lib/
python3.11/site-packages (from linearmodels) (0.4.3)
Requirement already satisfied: Cython>=0.29.34 in ./anaconda3/lib/pytho
n3.11/site-packages (from linearmodels) (3.0.6)
Requirement already satisfied: pyhdfe>=0.1 in ./anaconda3/lib/python3.1
1/site-packages (from linearmodels) (0.2.0)
Requirement already satisfied: formulaic>=0.6.5 in ./anaconda3/lib/pyth
on3.11/site-packages (from linearmodels) (0.6.6)
Requirement already satisfied: setuptools-scm[toml]<8.0.0,>=7.0.0 in ./
anaconda3/lib/python3.11/site-packages (from linearmodels) (7.1.0)
Requirement already satisfied: astor>=0.8 in ./anaconda3/lib/python3.1
1/site-packages (from formulaic>=0.6.5->linearmodels) (0.8.1)
Requirement already satisfied: interface-meta>=1.2.0 in ./anaconda3/li
b/python3.11/site-packages (from formulaic>=0.6.5->linearmodels) (1.3.
0)
Requirement already satisfied: typing-extensions>=4.2.0 in ./anaconda3/
lib/python3.11/site-packages (from formulaic>=0.6.5->linearmodels) (4.
7.1)
Requirement already satisfied: wrapt>=1.0 in ./anaconda3/lib/python3.1
1/site-packages (from formulaic>=0.6.5->linearmodels) (1.14.1)
Requirement already satisfied: python-dateutil>=2.8.1 in ./anaconda3/li
b/python3.11/site-packages (from pandas>=1.1.0->linearmodels) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in ./anaconda3/lib/python3.
11/site-packages (from pandas>=1.1.0->linearmodels) (2022.7)
Requirement already satisfied: packaging>=20.0 in ./anaconda3/lib/pytho
n3.11/site-packages (from setuptools-scm[toml]<8.0.0,>=7.0.0->linearmod
els) (23.0)
Requirement already satisfied: setuptools in ./anaconda3/lib/python3.1
1/site-packages (from setuptools-scm[toml]<8.0.0,>=7.0.0->linearmodels)
(68.0.0)
Requirement already satisfied: patsy>=0.5.2 in ./anaconda3/lib/python3.
11/site-packages (from statsmodels>=0.12.0->linearmodels) (0.5.3)
Requirement already satisfied: six in ./anaconda3/lib/python3.11/site-p
ackages (from patsy>=0.5.2->statsmodels>=0.12.0->linearmodels) (1.16.0)

```
In [3]: os.chdir('/Users/minpan/Desktop/Data Analysis  Data') # change working d
        d = pd.read_csv("panel-for-R (2).csv")
        d.head()
```

Out[3]:

formwt	oversamp	sample	panstat_2	panstat_3	mar1	...	wtpan12	wtpan123	wtpannr12	wtpa
1	1	9	1	1	5.0	...	0.414689	0.487828	0.435503	0.4
1	1	9	1	1	5.0	...	0.414689	0.487828	0.435503	0.4
1	1	9	1	1	1.0	...	0.414689	0.487828	0.435503	0.4
1	1	9	1	1	5.0	...	0.829377	0.858741	0.766632	0.8
1	1	9	1	1	5.0	...	0.829377	0.858741	0.766632	0.8

Variable of Interest

- abany: abortion if woman wants for any reason

Questions associated with this variable: Please tell me whether or not you think it should be possible for a pregnant woman to obtain a legal abortion if the woman wants it for any reason?

0:Inapplicable 1:Yes 2:No 8:Don't know 9:No answer

- partyid: Generally speaking, do you usually think of yourself as a Republican, Democrat, Independent, or what?

-99 .n: No answer -98 .d: Do not Know/Cannot Choose$ 0 Strong democrat 1 Not very strong democrat 2 Independent, close to democrat 3 Independent (neither, no response) 4 Independent, close to republican 5 Not very strong republican 6 Strong republican 7 Other party

- Degree: RESPONDENT'S DEGREE

0: Less than high school 1: High school 2: Associate/junior college 3: Bachelor's 4: Graduate

-97 .s: Skipped on Web -98 .d: Do not Know/Cannot Choose -99 .n: No answer

```
In [4]: #narrowing down to variables of interest
        sub=d[['abany','partyid','degree','idnum','panelwave']]
        sub=sub.dropna()
        sub.head()
```

Out[4]:

	abany	partyid	degree	idnum	panelwave
0	1.0	0.0	3.0	9	1
1	1.0	0.0	3.0	9	2
2	1.0	0.0	3.0	9	3
3	1.0	1.0	4.0	10	1
4	2.0	0.0	4.0	10	2

In [5]:
```python
# explore variables
summary = d['abany'].describe()
print(summary)
```

```
count    3189.000000
mean        1.592662
std         0.491416
min         1.000000
25%         1.000000
50%         2.000000
75%         2.000000
max         2.000000
Name: abany, dtype: float64
```

In [6]:
```python
summary2 = d['partyid'].describe()
print(summary2)
```

```
count    4784.000000
mean        2.823370
std         2.073284
min         0.000000
25%         1.000000
50%         3.000000
75%         5.000000
max         7.000000
Name: partyid, dtype: float64
```

In [7]:
```python
summary3 = d['degree'].describe()
print(summary3)
```

```
count    4811.000000
mean        1.648098
std         1.218126
min         0.000000
25%         1.000000
50%         1.000000
75%         3.000000
max         4.000000
Name: degree, dtype: float64
```

In [8]:
```python
# recode 'abany' to binary outcome (0 or 1), representing
# whether individuals support (1) or no support (0) abortion  if a woman

conditions = [
    (sub['abany'] == 1) ,
    (sub['abany'] == 2)]
choices = [1, 0]
sub['abn'] = np.select(conditions, choices, default=np.nan)
```

In [9]:
```python
# check the recoding result
pd.crosstab(index=sub["abn"], columns="count")
```

Out[9]:

col_0	count
abn	
0.0	1876
1.0	1296

In [10]:
```python
# recode 'partyid' to exlude other party Recode 'partyid' to exclude oth
# in this analysis I only examine democratic and republican affiliations

conditions = [
    (sub['partyid'] >= 0) & (sub['partyid'] <= 6)]
choices = [sub['partyid']]
sub['partyid1'] = np.select(conditions, choices, default=np.nan)

# check the recoding result
pd.crosstab(index=sub["partyid1"], columns="count")
```

Out[10]:

col_0	count
partyid1	
0.0	520
1.0	551
2.0	376
3.0	549
4.0	251
5.0	458
6.0	395

Naive ("pooled") OLS regression on the panel data

```python
lm_ols = smf.ols(formula = 'abn ~ partyid1 + degree', data = sub).fit()
print (lm_ols.summary())
```

```
                              OLS Regression Results
=================================================================================
======
Dep. Variable:                     abn    R-squared:
0.091
Model:                             OLS    Adj. R-squared:
0.091
Method:                  Least Squares    F-statistic:
155.2
Date:                 Tue, 12 Dec 2023    Prob (F-statistic):
5.89e-65
Time:                         20:19:07    Log-Likelihood:
-2042.7
No. Observations:                 3100    AIC:
4091.
Df Residuals:                     3097    BIC:
4110.
Df Model:                            2
Covariance Type:             nonrobust
=================================================================================
======
                   coef    std err          t      P>|t|      [0.025
0.975]
---------------------------------------------------------------------------------
-------
Intercept        0.4123      0.018     22.865      0.000       0.377
0.448
partyid1        -0.0536      0.004    -12.935      0.000      -0.062
-0.045
degree           0.0860      0.007     12.317      0.000       0.072
0.100
=================================================================================
======
Omnibus:                     19865.452    Durbin-Watson:
1.313
Prob(Omnibus):                   0.000    Jarque-Bera (JB):
362.062
Skew:                            0.374    Prob(JB):
2.39e-79
Kurtosis:                        1.502    Cond. No.
8.55
=================================================================================
======

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

Data Analysis:

I formulated hypotheses based on the expectation that individuals affiliated with the Democratic party are more inclined to support a woman's right to obtain an abortion for any reason, given the historical stance of the Democratic party on abortion rights. Additionally, I anticipated that individuals with higher educational degrees might be more likely to support abortion, assuming that those with higher education levels are more likely to endorse women's autonomy over their bodies.

The regression output shows that a one-unit increase in political affiliation from being Democratic to Republican is associated with a 5.4% decrease in the likelihood of a respondent supporting abortion if a woman wants it, all else being constant;

A one-unit increase in the respondent's educational degree is associated with an 8.6% increase in the likelihood of supporting abortion if a woman wants it, holding other variables constant, which indicates that higher educational levels are associated with increased support for abortion.

Both coefficients are statistically significant, indicated by p-values of 0.000.

The R-squared value (0.091) suggests that the model explains 9.1% of the variance in opinions

First Differences Regression

In [12]:
```
sub =sub.set_index(["idnum", "panelwave"])
sub
```

Out[12]:

idnum	panelwave	abany	partyid	degree	abn	partyid1
9	1	1.0	0.0	3.0	1.0	0.0
9	2	1.0	0.0	3.0	1.0	0.0
9	3	1.0	0.0	3.0	1.0	0.0
10	1	1.0	1.0	4.0	1.0	1.0
10	2	2.0	0.0	4.0	0.0	0.0
...	...	...	...	...	...	...
4500	1	2.0	6.0	1.0	0.0	6.0
4500	2	2.0	6.0	1.0	0.0	6.0
4500	3	2.0	6.0	1.0	0.0	6.0
4505	1	2.0	3.0	0.0	0.0	3.0
4505	2	2.0	5.0	0.0	0.0	5.0

3172 rows × 5 columns

```
In [16]: sub = sub.dropna()
         exog_vars = ["partyid1","degree"]
         exog = sub[exog_vars]
         lm_fd = FirstDifferenceOLS(sub.abn, exog)
         abn_fd = lm_fd.fit(cov_type='clustered', cluster_entity=True)
         print(abn_fd)
```

```
                    FirstDifferenceOLS Estimation Summary
===============================================================================
=========
Dep. Variable:                      abn   R-squared:
0.0026
Estimator:            FirstDifferenceOLS   R-squared (Between):
-0.2193
No. Observations:                  1757   R-squared (Within):
0.0007
Date:                  Tue, Dec 12 2023   R-squared (Overall):
-0.1952
Time:                          20:20:24   Log-likelihood
-1019.0
Cov. Estimator:               Clustered
                                          F-statistic:
2.3123
Entities:                          1330   P-value
0.0993
Avg Obs:                         2.3308   Distribution:
F(2,1755)
Min Obs:                         1.0000
Max Obs:                         3.0000   F-statistic (robust):
2.5361
                                          P-value
0.0795
Time periods:                         3   Distribution:
F(2,1755)
Avg Obs:                         1033.3
Min Obs:                         815.00
Max Obs:                         1286.0

                               Parameter Estimates
===============================================================================
=======
              Parameter   Std. Err.      T-stat    P-value    Lower CI      U
pper CI
-------------------------------------------------------------------------------
-------
partyid1        -0.0143      0.0084     -1.7142     0.0867     -0.0307
0.0021
degree          -0.0280      0.0193     -1.4493     0.1474     -0.0658
0.0099
===============================================================================
=======
```

Data Analysis:

The R-squared shows that there is only about 0.26% of the variance in the differenced 'abn' variable. This suggests that changes in 'partyid1' and 'degree' over time do not strongly predict changes in abortion support.

A one-unit shift from Democratic to Republican affiliation is associated with a 1.4% decrease in the likelihood of supporting abortion for any reason. This finding aligns with the my expectation of political affiliations influencing abortion views, but the effect size is small. The p-value for 'partyid1' is 0.0867, which is marginally above the conventional threshold of 0.05. This indicates that while there is some evidence of a relationship, it is not strong enough to be considered statistically significany.

Contrary to the OLS regression results and initial hypotheses, the coefficient for 'degree' suggests a negative association between changes in education and support for abortion. However, this relationship is not statistically significant given that p-value is 0.1474.

This result might suggest that within-individual changes in education level over time are not a strong predictor of changes in views on abortion, or it could reflect the influence of unobserved

In [14]:
```python
# Log-transformation of the variables
sub['log_abn'] = np.log(sub['abn'] + 1)
sub['log_partyid1'] = np.log(sub['partyid1'] + 1)
sub['log_degree'] = np.log(sub['degree'] + 1)
```

```
In [17]: exog_vars = ["log_partyid1","log_degree"] #e.g., exog_vars = ["hrs1", "e
         exog = sub[exog_vars]
         lm_fd = FirstDifferenceOLS(sub.log_abn, exog)
         abn_fd = lm_fd.fit(cov_type='clustered', cluster_entity=True)
         print(abn_fd)
```

```
                    FirstDifferenceOLS Estimation Summary
==============================================================================
=========
Dep. Variable:                    log_abn   R-squared:
0.0037
Estimator:            FirstDifferenceOLS   R-squared (Between):
-0.3169
No. Observations:                    1757   R-squared (Within):
0.0012
Date:                  Tue, Dec 12 2023   R-squared (Overall):
-0.2807
Time:                           20:20:31   Log-likelihood
-374.13
Cov. Estimator:                 Clustered
                                            F-statistic:
3.2237
Entities:                            1330   P-value
0.0400
Avg Obs:                           2.3308   Distribution:
F(2,1755)
Min Obs:                           1.0000
Max Obs:                           3.0000   F-statistic (robust):
2.9458
                                            P-value
0.0528
Time periods:                          3   Distribution:
F(2,1755)
Avg Obs:                           1033.3
Min Obs:                           815.00
Max Obs:                           1286.0

                            Parameter Estimates
==============================================================================
=========
              Parameter  Std. Err.     T-stat    P-value    Lower CI
Upper CI
------------------------------------------------------------------------------
----------
log_partyid1    -0.0358     0.0180    -1.9896     0.0468     -0.0711
-0.0005
log_degree      -0.0490     0.0343    -1.4286     0.1533     -0.1162
0.0183
==============================================================================
=========
```

Data Analysis:

The model reveals a statistically significant relationship between changes in political affiliation and the likelihood of supporting abortion. Specifically, a one-percentage point increase in 'partyid1', reflecting a shift towards Republican affiliation, is associated with a 3.5% decrease in the log-transformed likelihood of supporting abortion. This finding underscores the influential role of political ideology in shaping views on abortion, aligning with general expectations about political affiliations and social attitudes. This suggests that the effect of political affiliation changes on abortion views might be more multiplicative or exponential rather than linear.