Reproduction of ”Non-economic factors in violence: Evidence from organized crime, suicides and climate in Mexico”

Ming Senn Teo, Alica Modica, Bianca Rein Del Rosario, Seungmin Kim

April 26th, 2024

Abstract

Edward Miguel et al. (2018) examined the role of economic and non-economic factors in violence, using evidence from killings by drug trafficking organizations, homicides, and suicides in Mexico. Temperature was used as the main independent variable, given its well-established effect on both economic production and psychological factors that could be drivers of violence. Various controls were added to test their role in mediating the relationship between climate and violent outcomes. While finding only a partial effect of economic factors, non-economic factors were found to be significant in explaining the rise of violence. We have conducted a reproduction using the same statistical procedures and data as the original paper, excluding the analysis on DTO killings due to the unavailability of this data. We confirm the sign, magnitude, and statistical significance of the original estimates on violence.

1 Introduction

The paper written by Edward Miguel, Ceren Baysan, Marshall Burke, Felipe González and Solomon Hsiang (2018)[1] attempts to understand which factors lie behind the presence of violence, using evidence on killings by drug trafficking organizations, homicides, and suicides in Mexico. In particular, it investigates which ones between economic and non-economic factors better explain the relationship between temperature and violence. Temperature is used to model a ”taste for violence”, leveraging the well-established positive relationship between hot climates and spikes in violence. Temperature has been shown to cause not only impacts on economic production, but also psychological distress on individuals.

The conclusions of the study imply that temperature can induce large additional

increase in violence on homicides, and that this relationship cannot be explained by economic factors alone, as the limited impact of a cash transfer program suggests. Non-economic psychological and physiological factors that are affected by temperature, instead, likely play an important role in causing violence.
As suicide is know to be heavily influenced by mental distress, and the effect of temperature changes on the latter is well-established, this phenomenon is used in the study as a benchmark to show the role of psychological factors in temperature's effect on violence. The study ultimately confirms that inter-group and interpersonal violence respond to temperature in the same way suicide does.
The original paper also adds evidence to the already well-established relationship between suicides and climate, and proposes that, at least in the setting of the study, economic interventions to moderate the effect of climate change on rising intra-personal violence are not sufficient. We have conducted a reproduction using the same statistical procedures and data as the original paper. The data presents official records of homicides and suicides from Mexico's Bureau of Statistics (INEGI) for the period between January 1990 and December 2010. In order to avoid confounding with the Mexican Drug War, the data is split into two periods, "pre-war" (January 1990 - December 2006) and "war" (January 2007 - December 2010), and only the first one is used in the analysis of homicides and suicides.
We have excluded from our reproduction the tables and regressions using DTO (Drug Trafficking Organizations) killings as a variable representing intergroup violence, since it was not present in the dataset available for reproduction.

2 Reproducibility

There weren't any coding errors found from the original STATA code by the authors. However, when trying to reproduce the results, the data for drug traffic organizations(DTO) killings was not included in the public datasets. As a result, we were not able to reproduce figures and tables that involved this variable.

2.1 Challenges

2.1.1 Figure 1: We did not produce this figure as this is just a visual that shows the spectrum of violence and there was no relevant STATA code.

2.1.2 Figure 2: We were able to reproduce it almost identically, except for the DTO killings data and the format of the X-axis labels due to Python.

2.1.3 Figure 3, 5, 6: Although the trends of the graphs are similar, the estimates of our parameters and confidence intervals are different than in the paper.

The regressions for these figures include three fixed effects: municipality, year, and month. However, the package PanelOLS from linearmodels cannot take in more than two fixed effects so we could not properly account for the time-fixed effects[2]. Additionally, we tried to calculate fixed effects using dummy variables along with the statsmodel.api but the results were drastically different from the paper and the runtime was 45 minutes[A-1, A-2]. As a result, we were only able to calculate with municipality and a combined month and year fixed effects so the reproduced figure was not the same as the one in the original paper.

Though the values were different, the trends were the same so it does not affect the main argument of the paper. Figure 3 still shows a positive linear relationship between temperature and homicide rates. Similarly, Figure 5 still shows a positive linear relationship between temperature and suicide rates. The regression estimates in figure 6 are similar to the paper with only a difference in confidence intervals.

One minor discrepancy we saw that wasn't necessarily a coding error for figures 3 and 5 but perhaps the way STATA produces visualizations is with the widths of the temperature bins. In the original figure, the bins are a width of 3 but we believe they just failed to specify that the left point was inclusive and the endpoint was exclusive.

2.1.4 Figure 4:

This figure was not reproduced as the results came from another paper and the panel from the paper that was presented as evidence for the claim is seen in other figures and tables.

2.1.5 Table 1:

The standard deviations within statistics, which is "the standard deviation of the corresponding variable after removing municipality fixed effects", were not calculated for this table as the package to calculate this, f_oneway from scipy.stats, cannot calculate the statistics for more than 9 groups and there are over 2000 groups(municipalities). The original STATA code uses loneway which is similar as they both conduct a one-way ANOVA test.

2.1.6 Table 2:

Referring to the notes that columns 2-6 used data in the periods from 2007 to 2010, the data used on the reproduction was actually before 2007. With these contradictory conditions of data, the figures in the result differ. However, the overall trends look consistent in that each figure has identical signs and is close in magnitude.

2.1.7 Table 3:

With the original stata code, 'areg', it was not possible to calculate the means of the dependent variable and 'within standard deviations.' However, the coefficients of dependent variables, numbers of observations, and R-squared are the same. Standard deviations and p-values are found to be different from the original table.

2.1.8 Table 4, 5, 6, 8 The regressions match those in the original paper, with only some minor differences in the coefficients, which will be discussed in the next section.

2.1.9 Table 7: This regression was not performed as the dependent variable, DTO, was not present in the dataset.

3 Regression model

The regression estimates and confidence intervals derived from figures 3 and 5:

$$\mathrm{y}_{nsmt} = \beta \mathrm{Temp}_{nsd} + \delta \mathrm{Precip}_{nsd} + \xi_d + \lambda_n + \zeta_n + \epsilon_{nsd} \tag{1}$$

where y is the number of DTO killings, homicides, or suicides per 100,000 inhabitants in municipality n, state s, month m, and year t; ξ and λ are full sets of month and year fixed effects; ζ is a full set of municipality fixed effects, respectively; Temp is average temperature, measured in degrees celsius; Precip is total precipitation, measured in thousands of millimeters; and ϵ is an error term clustered at the state level. The x-axis indicates the deviation from the municipality temperature average while the y-axis is the deviation from the municipality average for the variable in study(homicide and suicide rates).

The regression estimates and confidence intervals derived from figure 6:

$$\mathrm{y}_{smt} = \alpha + \xi_m + \lambda_t + \zeta_s + \sum_{k=-6}^{6} \beta_{t+k} \mathrm{Temp}_{sm,t+k} + \sum_{k=-6}^{6} \delta_{t+k} \mathrm{Precip}_{sm,t+k} + \epsilon_{smt} \tag{2}$$

where y is the number of DTO killings, homicides, or suicides per 100,000 inhabitants in municipality s, month m, and year t; ξ_m, λ_t, and ζ_s are month, year, and municipality fixed effects; $Temp_{smt}$ and $Precip_{smt}$ are temperature measured in degrees celsius and precipitation measured in thousands of millimeters respectively. ϵ is an error term clustered at the state level. The x-axis indicates the lags and leads of the effects of temperature in month t -1 (lags) and month t + 1 (leads) on violence in month t, while the y-axis is the deviation from the municipality average for the variable in study(homicide and suicide rates).

For table 1, it was finding mean and standard deviation that no other special regression model was required.

For table 2, column(1), which is column 4 in the original paper, required deriving regression result with the effect of temperature and precipitation and no other fixed effects considered. The regression expression was therefore the following:

$$\text{Homicides} = \beta_0 + \beta_1 \text{Temperature} + \beta_2 \text{Precipitation} + \epsilon \tag{3}$$

The column(2), which is column(6) in the original paper, required the model to be modified with fixed effects of Time. By adding 'TimeEffects' into the code, we made the regression expression like the following where the greek letter xi stands for the Time fixed effect:

$$\text{Homicides} = \beta_0 + \beta_1 \text{Temperature} + \beta_2 \text{Precipitation} + \xi + \epsilon \tag{4}$$

Lastly, the column(3), originally column(5), was with the fixed effects of both Time effects and Entity effects. Same as the column(2), by adding 'TimeEffedts' and 'EntityEffects,' the regression expression for the column 3 was the following, where the greek letter epsilon stands for entity effect:

$$\text{Homicides} = \beta_0 + \beta_1 \text{Temperature} + \beta_2 \text{Precipitation} + \xi + \lambda + \epsilon \tag{5}$$

For table 3, there were 4 regressions for total, where independent variables and fixed effects are all same but dependent variables changing by the regressions. The dependent variable for the first regression was homicide rate that the equation was following:

$$\text{Homicides} = \beta_0 + \beta_1 \text{Temperature} + \beta_2 \text{Precipitation} + \text{C}_{month} + \text{C}_{year} + \text{C}_{id} + \epsilon \tag{6}$$

Same logic as the first regression, the dependent variable for the second regression was 'car thefts,' 'extortion' for the third and 'kidnap' for the third. Modifying only dependent variable, the regression is like the followings:

$$\text{Car Theft} = \beta_0 + \beta_1 \text{Temperature} + \beta_2 \text{Precipitation} + \text{C}_{month} + \text{C}_{year} + \text{C}_{id} + \epsilon \tag{7}$$

$$\text{Extortion} = \beta_0 + \beta_1 \text{Temperature} + \beta_2 \text{Precipitation} + \text{C}_{month} + \text{C}_{year} + \text{C}_{id} + \epsilon \tag{8}$$

$$\text{Kidnap} = \beta_0 + \beta_1 \text{Temperature} + \beta_2 \text{Precipitation} + \text{C}_{month} + \text{C}_{year} + \text{C}_{id} + \epsilon \tag{9}$$

In order to test whether municipality-level economic variables mediate the temperature-violence relationship, the original paper presents different regressions controlling for municipality-level income variables. The researchers estimated in STATA a panel data regression with month and year fixed effects, as well as municipality fixed effects. Year and month effects were added by creating dummies for each unique year and month, while entity effects where added using specific STATA syntex. In Python, panel data regression is not as straightforward as in STATA. While creating dummies for each year, month and municipality would . We set on the use of the package *PanelOLS* from the package *linearmodels*. This package allows to add entity and time fixed effects, but can support only one variable for each effect. Therefore, we created a variable Date to capture the effect of both the year variable and the month variables, and used it to add time fixed effects to the regression. The unique code identifier for each municipality was used to capture entity effects. Moreover, as in STATA, we weighted the analysis on the variable *popw* and we clustered the standard errors at the state level. The regression equation estimated in Table 4 is:

$$\text{Homicides}_{nsmt} = \beta\text{Temp}_{nsd} + \delta\text{Precip}_{nsd} + \xi_d + \lambda_n + \gamma\text{Temp}_{nsd} \times \text{Control}_{nsd} + \epsilon_{nsd} \quad (10)$$

where y is the number of homicides for 100,000 people in municipality n on day d; ξ is a set of time fixed effects; λ is a set of municipality fixed effects; Temp is average temperature measured in degrees Celsius; Precip is total precipitation measured in thousands of millimeters; and ε is an error term clustered at the state level. The results in Table 4 match those in the original paper. The effect of temperature on the homicide rate is positive and statistically significant at the 1% confidence level. It appears that both income and the Gini Index, respectively measures of income and of economic inequality, do not mediate the relationship between temperature and violence, as shown by small and non significant coefficients on these variables. The same conclusions can be drawn for the variables representing average temperatures and air conditioning. There is only little evidence that the latter could induce a reduction in violence in richer households that have air conditioning, with a small coefficient with significance at the 10% confidence level.

A second approach to studying the role of economic factors on violence was brought forward by the researchers using data on a large-scale cash transfer program, Progresa. Transfers to families started in 1998, therefore the data is filtered to exclude years when the program had not been established yet.

$$\text{Homicides}_{nsmt} = \beta\text{Temp}_{nsd} + \delta\text{Precip}_{nsd} + \sigma\text{Progresa}_{nsd} + \gamma\text{Temp}_{nsd} \times \text{Progresa}_{nsd} + \xi_d + \lambda_n + \epsilon_{nsd} \quad (11)$$

The effect of temperature and Progresa cash transfers on homicides is estimated

by adding, as in the previous regressions, day and municipality fixed effects, as well as an interaction γ between the two main regressors.
The results of the reproduction confirmed the outcomes of the original paper, with only minor differences of 0.001 in all but one coefficient. We believe that this difference can be discarded, as it could be caused simply by the use of a different software. The results of both the original and the reproduced regressions conclude that the effect of these variables is small and not statistically significant.

The third and last approach to studying the role of economic factors on violence was to explore the impacts on violence during economically critical times. The following regression model is estimated, using growing season, households in rural areas, and workers in the agricultural sector as controls:

$$\text{Homicides}_{nsmt} = \beta\text{Temp}_{nsd} + \delta\text{Precip}_{nsd} + \sigma\text{Control}_{nsd} + \xi_d + \lambda_n \tag{12}$$

The results we gained show that temperature shocks during the growing season, the percentage of households living in rural areas, and the percentage of workers in the agricultural sector fail to explain the relationship between climate and violence. All of the coefficients are small and not statistically significant. There is only one minor difference from the original paper that can be spotted, which is the fact that the Temperature variable when controlling for households living in rural areas has a higher level of significance at 1% confidence level.

Lastly, the regressions in Table 8 in the original paper (Table 7 in this paper), investigate the role of psychological factors in the link between temperature and violence. The following regression model is estimated, using different economic variables as controls:

$$\text{Suicides}_{nsmt} = \beta\text{Temp}_{nsd} + \delta\text{Precip}_{nsd} + \sigma\text{Control}_{nsd} + \gamma\text{Temp}_{nsd} \times \text{Control}_{nsd} + \xi_d + \lambda_n + \epsilon_{nsd} \tag{13}$$

The results we obtained confirm the fact that the response of suicides to temperature mirrors that of homicides, and that economic variables fail to explain the relationship between variations in climate and violence. Most coefficients are small and not significant, and the only ones with significance have signs opposite to what would be expected. Only some minor differences from the original paper can be spotted, such as the fact that, in column (6), the interaction variable between Temperature and Progresa Transfers has a coefficient of -0.002 instead of -0.001. Additionally, in column (7), Python, unlike STATA, absorbed the variable *growing_season* as it had the same effect as the interaction between it and the

Temperature variable, also included in the regression.

4 Conclusion

Overall, the reproduction was possible but challenging because of the amount of data that we had to process as well as finding compatible Python packages that can replicate STATA functions. PanelOLS and statsmodel proved to be helpful for running the regressions but had computational limitations such as the number of fixed effects and runtime. Additionally, chatGPT helped find packages to use and gave skeleton code.

Despite these challenges, the reproduced evidence is closely aligned with the original paper's three main findings: supporting the established relationship between temperature and violence(intergroup and interpersonal), the lack of influence of economic factors on this association, and the presence of non-economic factors for explaining this link. As mentioned, the main difference between the reproduction and the original paper was the absence of the DTO killings variable to strengthen various analyses like proving the relationship between temperature and intergroup violence. Despite this, other variables like homicide rate are still strong metrics for evidence for these claims.

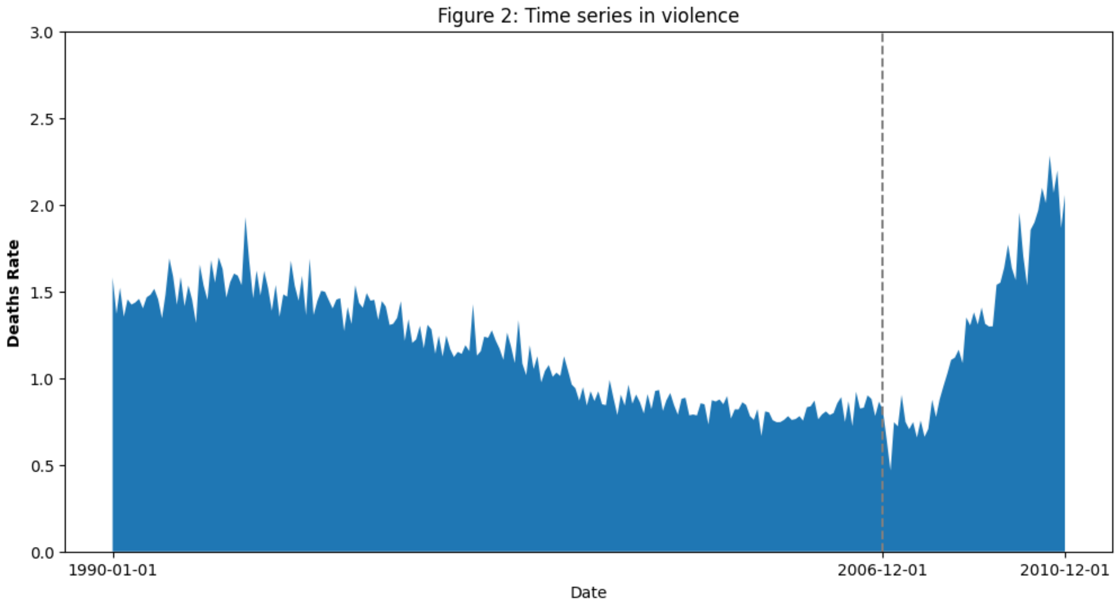
References

[1] Felipe González Solomon Hsiang Ceren Baysan, Marshall Burke and Edward Miguel. Economic and Non-Economic Factors in Violence: Evidence from Organized Crime, Suicides and Climate in Mexico. *National Bureau of Economic Research*, pages 1–40, 2018.

[2] Matheus Facure Alves. Causal inference for the brave and true, 2022. Last accessed 26 April 2024.

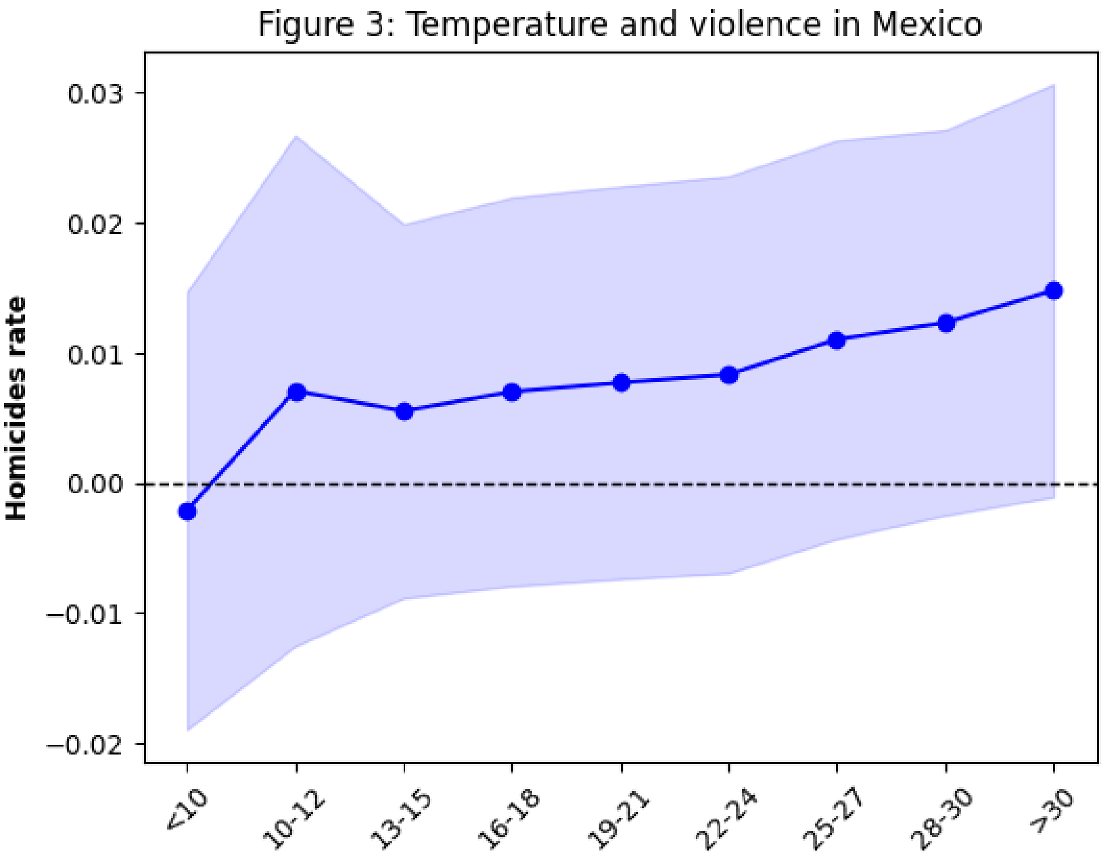
5 Figures

Figure 2: Time Series in Violence



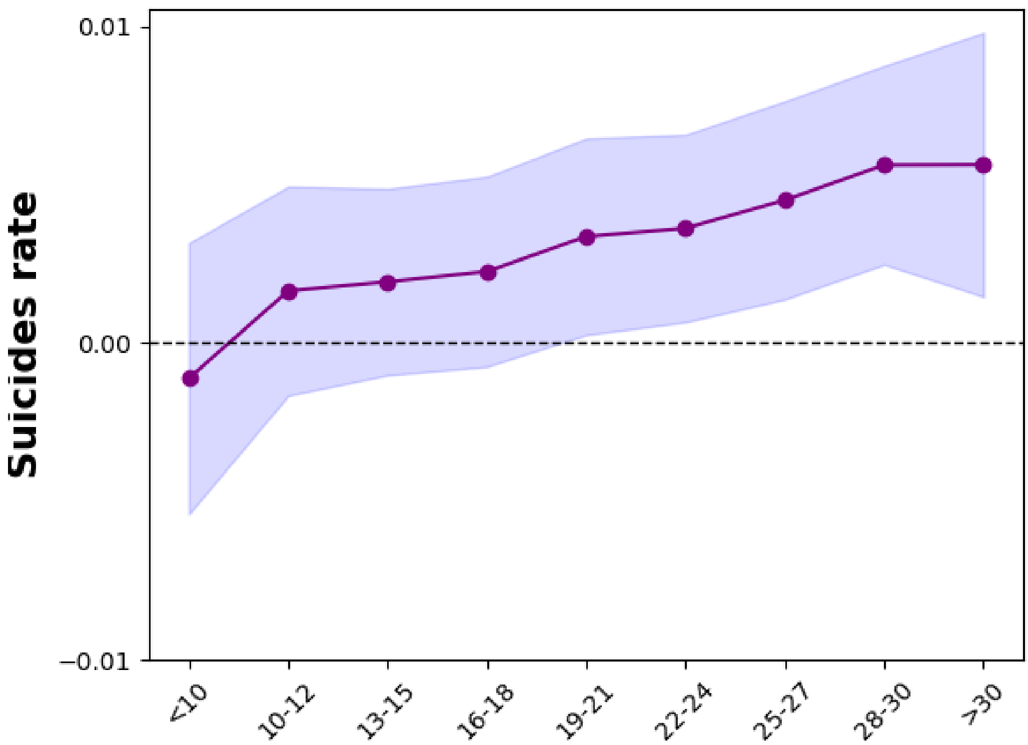
Notes: Time averages (weighted by population) for homicide rate in all municipalities in Mexico. The dashed vertical black line denotes the beginning of the Mexican Drug War.

Figure 3: Temperatures and Violence in Mexico



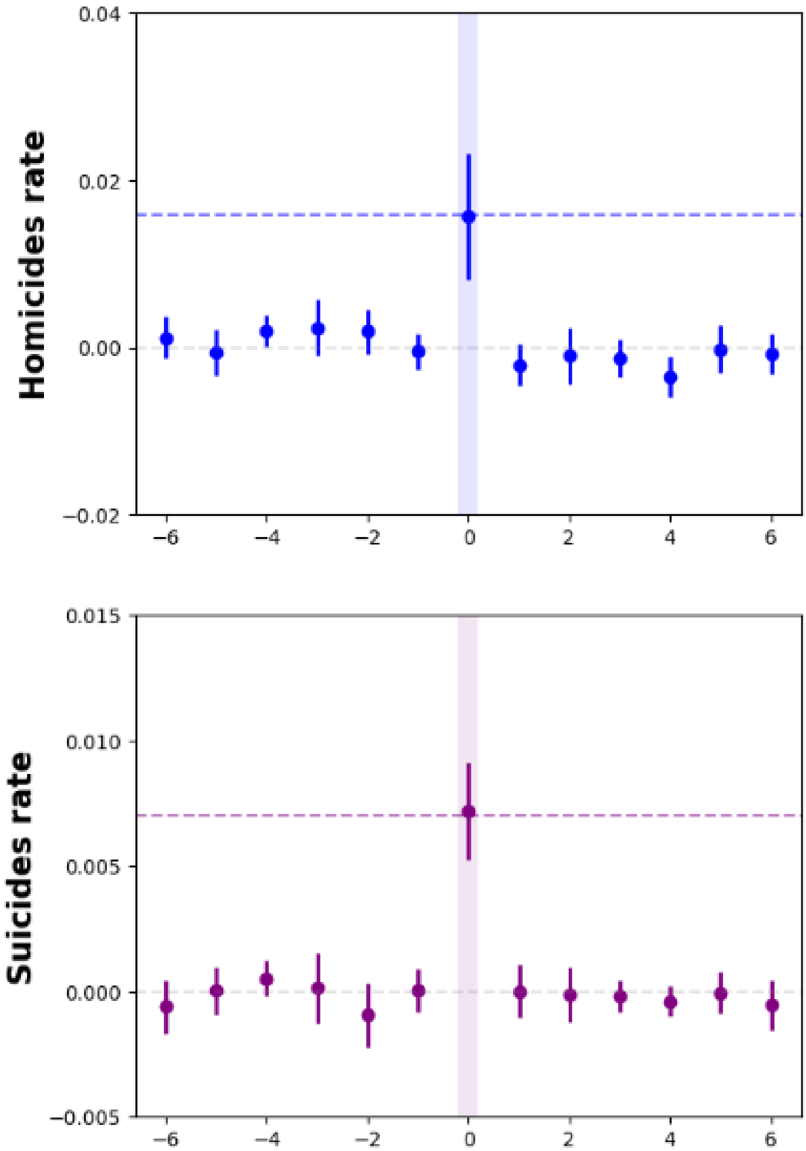
Notes: Temperature response functions for suicides using temperature bins of width 2°C. Regression estimates and confidence intervals derived from Equation (1).

Figure 5: Temperatures and suicides



Notes: Temperature response functions for suicides using temperature bins of width 2°C. Regression estimates and confidence intervals derived from Equation (1).

Figure 6: Temporal distribution of estimates



Notes: Regression estimates of β_{t+k} of equation 2 from section 3 Regression.

6 Tables

Table 1. Descriptive statistics

Period:	*January 1990 - December 2006*			*January 2007 - December 2010*		
	Mean	St. Dev.	St. Dev within	Mean	St. Dev.	St. Dev within
Homicides per 100,000 inhabitants	0.98	5.23	-	0.83	4.13	-
Suicides per 100,000 inhabitants	0.21	1.92	-	0.26	2.21	-
Population	39057.12	116901.42	-	44584.62	130760.44	-
Temperature (°C)	20.10	5.00	-	20.05	5.09	-
Precipitation (millimeters)	93	112	-	81	107	-
Municipalities	2,456			2,456		
Observations	494,724			117,696		

Note: Standard errors clustered at the state level. Standardized effects in brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 2. Temperature and violence in Mexico

Dependent variable:	*Homicides*		
	(1)	(2)	(3)
Temperature	0.0729***	0.0218	0.033
	(0.000)	(0.317)	(0.114)
	[0.017]	[0.022]	[0.021]
Precipitation	-0.2087*	-0.2151*	-0.0078
	(0.089)	(0.082)	(0.799)
	[0.123]	[0.124]	[0.031]
Municipality F.E.	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes
Month F.E.	Yes	No	Yes
Month-state F.E.	No	No	Yes
State trends	No	Yes	No
Observations	117,458	117,458	117,458

Note: Standard errors clustered at the state level. Standardized effects in brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 3. Temperature and economically motivated crimes

Dependent variable:	*Homicides*	*Car thefts*	*Extortions*	*Kidnappings*
	(1)	(2)	(3)	(4)
Temperature	0.050***	0.067	-0.005	-0.007
	(0.000)	(0.441)	(0.323)	(0.635)
	[0.013]	[0.086]	[0.006]	[0.001]
Precipitation	-0.285	-0.363	0.220	0.060
	(0.571)	(0.916)	(0.317)	(0.268)
	[0.503]	[3.429]	[0.220]	[0.054]
Mean of dep. variable	-	-	-	-
(Within st. dev.)	(-)	(-)	(-)	(-)
Municipality, year & month F.E.	Yes	Yes	Yes	Yes
Observations	1535	1535	1535	1535
R^2	0.714	0.886	0.603	0.392

Note: Standard errors clustered at the state level. Standardized effects in brackets. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 4. Interaction with economic variables

Dependent variable:	*Homicides*			
	(1)	(2)	(3)	(4)
Temperature	0.014***	0.016***	0.022***	0.016***
	(0.004)	(0.004)	(0.005)	(0.004)
× **Income (1990)**	0.003			
	(0.002)			
× **Gini (1990)**		0.001		
		(0.003)		
× **Houses with air conditioning (2010)**			-0.002*	
			(0.001)	
× **Average temperature (1990-2010)**				0.000
				(0.002)
Municipality F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
Month F.E.	Yes	Yes	Yes	Yes
Observations	486,132	482,868	121,056	493,908

Note: Standard errors clustered at the state level. *p<0.1; **p<0.05; ***p<0.01

Table 5. Progresa transfers

Dependent variable:	*Homicides*		
	(1)	(2)	(3)
Temperature	0.013^{***}	0.013^{***}	0.013^{***}
	(0.003)	(0.003)	(0.003)
Progresa Transfers		0.004	0.004
		(0.011)	(0.012)
Progresa Transfers x Temperature			-0.002
			(0.002)
Municipality F.E.	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes
Month F.E.	Yes	Yes	Yes
Observations	262,992	262,992	262,992
Note: Standard errors clustered at the state level.	*p<0.1; **p<0.05; ***p<0.01		

Table 6. Interaction with agricultural variables

Dependent variable	*Homicides*			
	(1)	(2)	(3)	(4)
Temperature	0.016***	0.013***	0.015***	0.015***
	(0.004)	(0.005)	(0.005)	(0.005)
× **Growing season indicator**		0.005		
		(0.006)		
× **Households in rural areas (1990)**			-0.000	
			(0.002)	
× **Workers in agricultural sector (1990)**				-0.000
				(0.002)
Municipality F.E.	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes
Month F.E.	Yes	Yes	Yes	Yes
Observations	493,908	493,908	488,784	488,784

Note: Standard errors clustered at the state level. *p<0.1; **p<0.05; ***p<0.01

Table 7. Temperature and suicides in Mexico

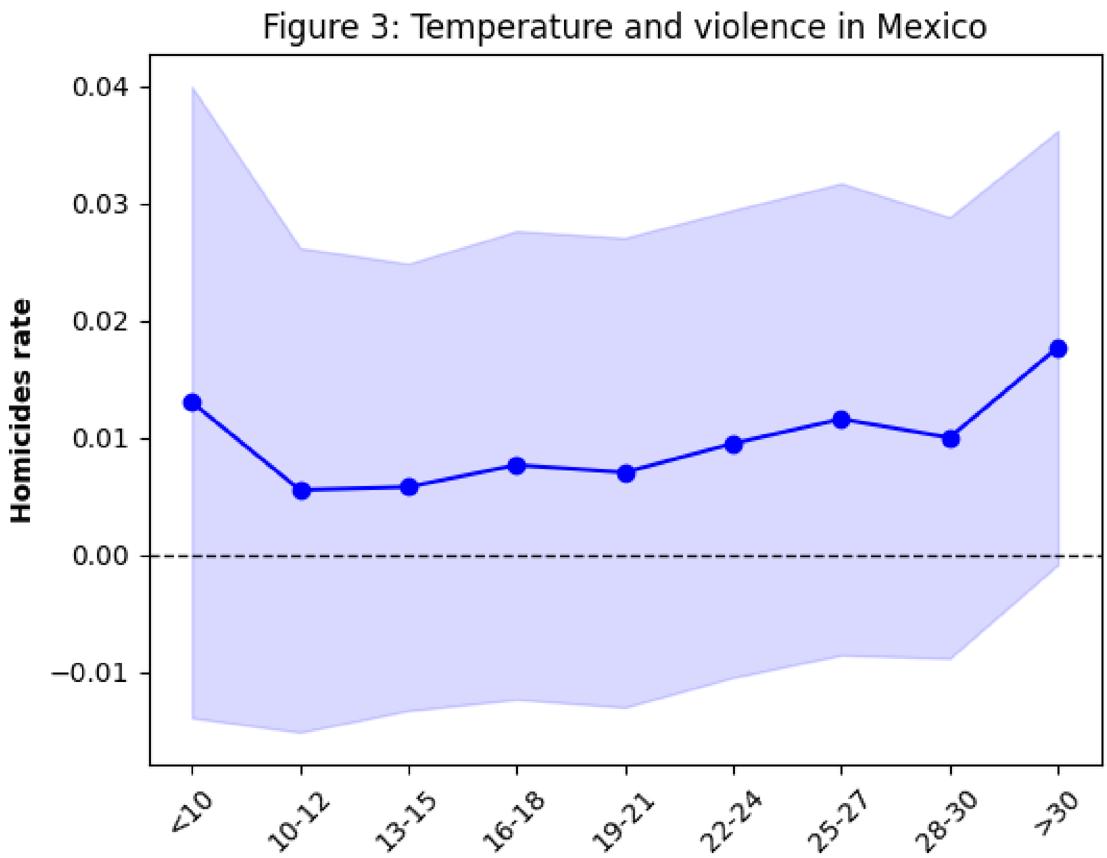
Dependent variable is suicides rate per 100,000 inhabitants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Temperature	0.007***	0.006***	0.007***	0.005***	0.007***	0.009***	0.009***	0.007***	0.006***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
× **Income**		0.001							
		(0.001)							
× **Gini**			0.000						
			(0.001)						
× **Homes with Air Conditioning**				0.001					
				(0.001)					
× **Average Temperature**					0.001*				
					(0.001)				
× **Temperature × Progresa**						-0.002			
						(0.002)			
× **Progresa Transfers**						-0.009*			
						(0.005)			
× **Growing Season**							-0.004**		
							(0.002)		
× **Households in Rural Areas**								-0.000	
								(0.001)	
× **Agricultural Workers**									-0.000
									(0.001)
Municipality F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	493,908	486,132	482,868	121,056	493,908	262,992	493,908	488,784	488,784

*p<0.1; **p<0.05; ***p<0.01

7 Appendix

A-1: Figure 3 - Attempt



Notes: Figure with a manual calculation of fixed effects and using the package statsmodel.

A-2: Figure 3 - Attempt Code

```
m = sm.regression.linear_model.OLS(y, X, weights=fig_3_1['popw'])
r = m.fit(cov_type='cluster', cov_kwds={'groups': fig_3_1['state']})
[12]                                                                    Python
...
/Library/Frameworks/Python.framework/Versions/3.8/lib/python3.8/site-packages/statsmodels/regression/linear_model.py:922: ValueWarning: Weights are not
supported in OLS and will be ignoredAn exception will be raised in the next version.
  warnings.warn(msg, ValueWarning)
/Library/Frameworks/Python.framework/Versions/3.8/lib/python3.8/site-packages/statsmodels/base/model.py:130: ValueWarning: unknown kwargs ['weights']
  warnings.warn(msg, ValueWarning)

c = r.params
ci = r.conf_int()
rdf = pd.DataFrame({'coefficient': c, 'CI_lower': ci[0], 'CI_upper': ci[1]})
rdf = rdf.iloc[2:11]
[27]                                                                    Python
```

Notes: Dummy variables for fixed effects made with pandas and regression calculated with statsmodel.

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from linearmodels import PanelOLS
import datetime
from scipy.stats import f_oneway
import statsmodels.api as sm
```

Data Cleaning

Tables

In [2]:
```
muni = pd.read_stata('mexico_muni_jebo.dta')
muni.head()
```

Out[2]:

	muni_code	state	municipality	year	month	temp_pop	precip_pop	id	
0	1001	Aguascalientes	Aguascalientes	2001.0	4.0	23.817400	0.113894	1.0	0
1	1001	Aguascalientes	Aguascalientes	1998.0	4.0	22.207600	0.000000	1.0	0
2	1001	Aguascalientes	Aguascalientes	2001.0	11.0	16.335100	0.038703	1.0	0
3	1001	Aguascalientes	Aguascalientes	2000.0	2.0	17.702499	0.002461	1.0	0
4	1001	Aguascalientes	Aguascalientes	2001.0	7.0	21.856300	1.110245	1.0	0

5 rows × 57 columns

In [3]:
```
# Add a date column to support datetime data type
muni['day'] = 1
muni['date'] = pd.to_datetime(dict(year=muni.year, month=muni.month, day=muni.day))
# Convert Year and Month to Integers
muni['year'] = muni['year'].astype('int')
muni['month'] = muni['month'].astype('int')
```

In [4]:
```
state = pd.read_stata("mexico_state_jebo.dta")
state.head()
```

```
C:\Users\blahb\anaconda3\lib\site-packages\pandas\io\stata.py:1457: UnicodeWarning:
One or more strings in the dta file could not be decoded using utf-8, and
so the fallback encoding of latin-1 is being used.  This can happen when a file
has been incorrectly encoded by Stata or some other software. You should verify
the string values returned are correct.
  warnings.warn(msg, UnicodeWarning)
```

Out[4]:

	id	year	state	month	temperature	homsegob_rate	kid_rate	ext_rate
0	1804.0	1990.0	Aguascalientes	1.0	15.166670	NaN	NaN	NaN
1	1804.0	1990.0	Aguascalientes	2.0	14.666670	NaN	NaN	NaN
2	1804.0	1990.0	Aguascalientes	3.0	17.866671	NaN	NaN	NaN
3	1804.0	1990.0	Aguascalientes	4.0	19.833330	NaN	NaN	NaN
4	1804.0	1990.0	Aguascalientes	5.0	24.600000	NaN	NaN	NaN

In [5]:

```
# Add a date column to support datetime data type
state['day'] = 1
state['date'] = pd.to_datetime(dict(year=state.year, month=state.month, day=state.d
# Convert Year and Month to Integers
state['year'] = state['year'].astype('int')
state['month'] = state['month'].astype('int')
```

Figure 2: Time Series in Violence

In [6]:

```
# Aggregate the Homicide Rate by date and t column weighting it to the population
fig2_df = muni.groupby(['t', 'date']).agg({'hom_rate': lambda x: (x * muni['popw'])
fig2_df.head()
```

Out[6]:

	t	date	hom_rate
0	1.0	1990-01-01	1.579638
1	2.0	1990-02-01	1.369551
2	3.0	1990-03-01	1.520228
3	4.0	1990-04-01	1.353619
4	5.0	1990-05-01	1.453345

In [7]:

```
# Create the plot
fig, ax = plt.subplots(figsize = (12,6))

# Plot ranges and points
plt.stackplot(fig2_df['date'], fig2_df['hom_rate'])
plt.axvline(x = datetime.date(2006, 12, 1), linestyle = "--", color = 'gray')

# Styling
plt.xlabel('Date')
plt.ylabel('Deaths Rate', fontweight='bold')
plt.title('Figure 2: Time series in violence')
plt.ylim(0, 3)
ticks = fig2_df['date'].iloc[[0, 203,251]]
labels = [ i.strftime("%Y-%m-%d") for i in ticks ]
ax.set_xticks(ticks)
ax.set_xticklabels(labels)
```

Out[7]:
```
[Text(7305.0, 0, '1990-01-01'),
 Text(13483.0, 0, '2006-12-01'),
 Text(14944.0, 0, '2010-12-01')]
```

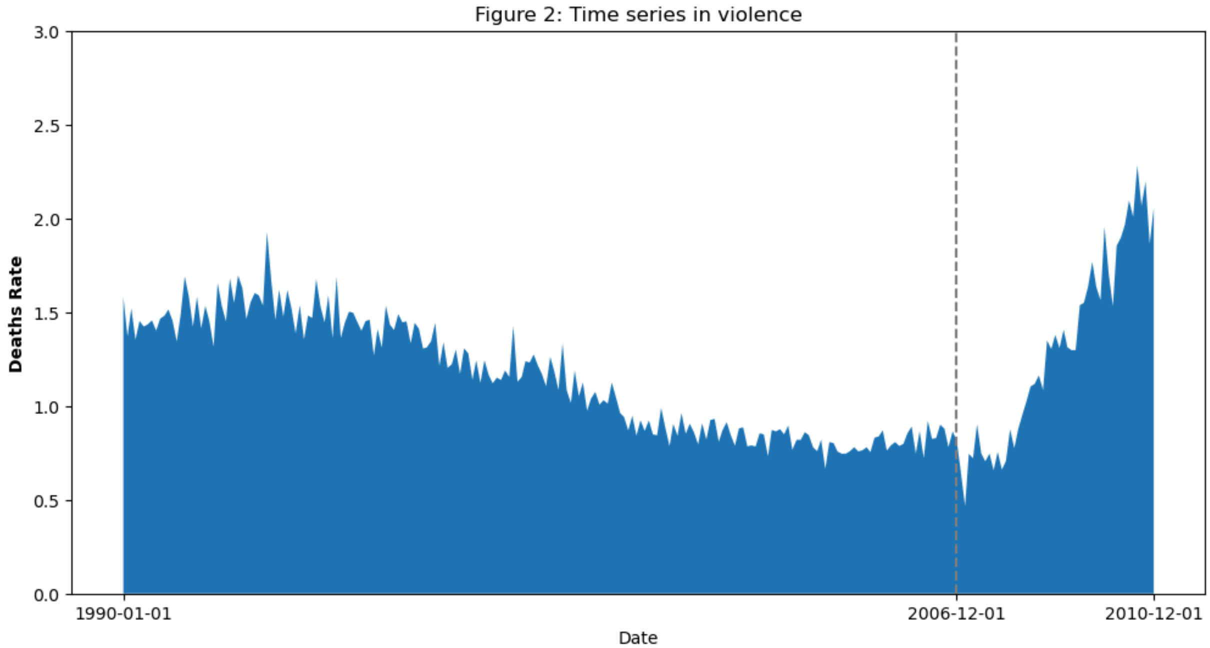


Figure 3: Temperature and violence in Mexico

In [8]:
```
# Bins for the temperature
t_bins = ['temp_bin_pop_0_9', 'temp_bin_pop_10_12', 'temp_bin_pop_13_15', 'temp_bin

# More Table Data cleaning: Filter for years and drop Nan rows
fig_3 = muni[muni['year'] < 2007].dropna(subset = ['hom_rate', 'precip_pop'] + t_bi

# Set the Index for the Entity Effects and Time Effects
fig_3 = fig_3.set_index(['muni_code', 'date'])
fig_3 = fig_3[['state', 'hom_rate', 'temp_pop','precip_pop', 'popw'] + t_bins]
fig_3.head()
```

Out[8]:

muni_code	date	state	hom_rate	temp_pop	precip_pop	popw	temp_bin
1001	2001-04-01	Aguascalientes	0.301511	23.817400	0.113894	655030.9375	
	1998-04-01	Aguascalientes	0.642868	22.207600	0.000000	655030.9375	
	2001-11-01	Aguascalientes	0.297347	16.335100	0.038703	655030.9375	
	2000-02-01	Aguascalientes	0.000000	17.702499	0.002461	655030.9375	
	2001-07-01	Aguascalientes	0.149856	21.856300	1.110245	655030.9375	

In [9]:

```
# Regressions with Time and Municipality Fixed Effects that are weighted population
fig3_mod = PanelOLS.from_formula('hom_rate ~ ' + ' + '.join(t_bins) + '+ precip_pop
fig3_results = fig3_mod.fit(cov_type='clustered', clusters = fig_3['state'])
fig3_coef = fig3_results.params
# 95% confidence level
fig3_ci = fig3_results.conf_int(level =  0.95)

fig3_result_df = pd.DataFrame({'coefficient': fig3_coef, 'CI_lower': fig3_ci['lower
fig3_result_df
```

Out[9]:

	coefficient	CI_lower	CI_upper
temp_bin_pop_0_9	-0.002133	-0.018948	0.014681
temp_bin_pop_10_12	0.007094	-0.012527	0.026716
temp_bin_pop_13_15	0.005532	-0.008829	0.019894
temp_bin_pop_16_18	0.007021	-0.007913	0.021954
temp_bin_pop_19_21	0.007710	-0.007377	0.022796
temp_bin_pop_22_24	0.008335	-0.006913	0.023583
temp_bin_pop_25_27	0.011026	-0.004287	0.026338
temp_bin_pop_28_30	0.012324	-0.002477	0.027124
temp_bin_pop_31_Inf	0.014794	-0.001063	0.030650

In [10]:

```
# Create the plot
fig, ax = plt.subplots()

# Plot ranges and points
ax.plot(fig3_result_df.index, fig3_result_df['coefficient'], '-o', color='blue')
ax.fill_between(fig3_result_df.index, fig3_result_df['CI_lower'], fig3_result_df['C
```

```
# Styling
ax.set_title('Figure 3: Temperature and violence in Mexico')
ax.axhline(0, color='black', linewidth=1, linestyle='dashed')
ax.set_xticks(fig3_result_df.index)
ax.set_xticklabels(["<10", "10-12", "13-15", "16-18", "19-21", "22-24", "25-27", "2
ax.set_ylabel('Homicides rate', fontweight='bold')
ax.set_xlabel('')

# Show the plot
plt.show()
```

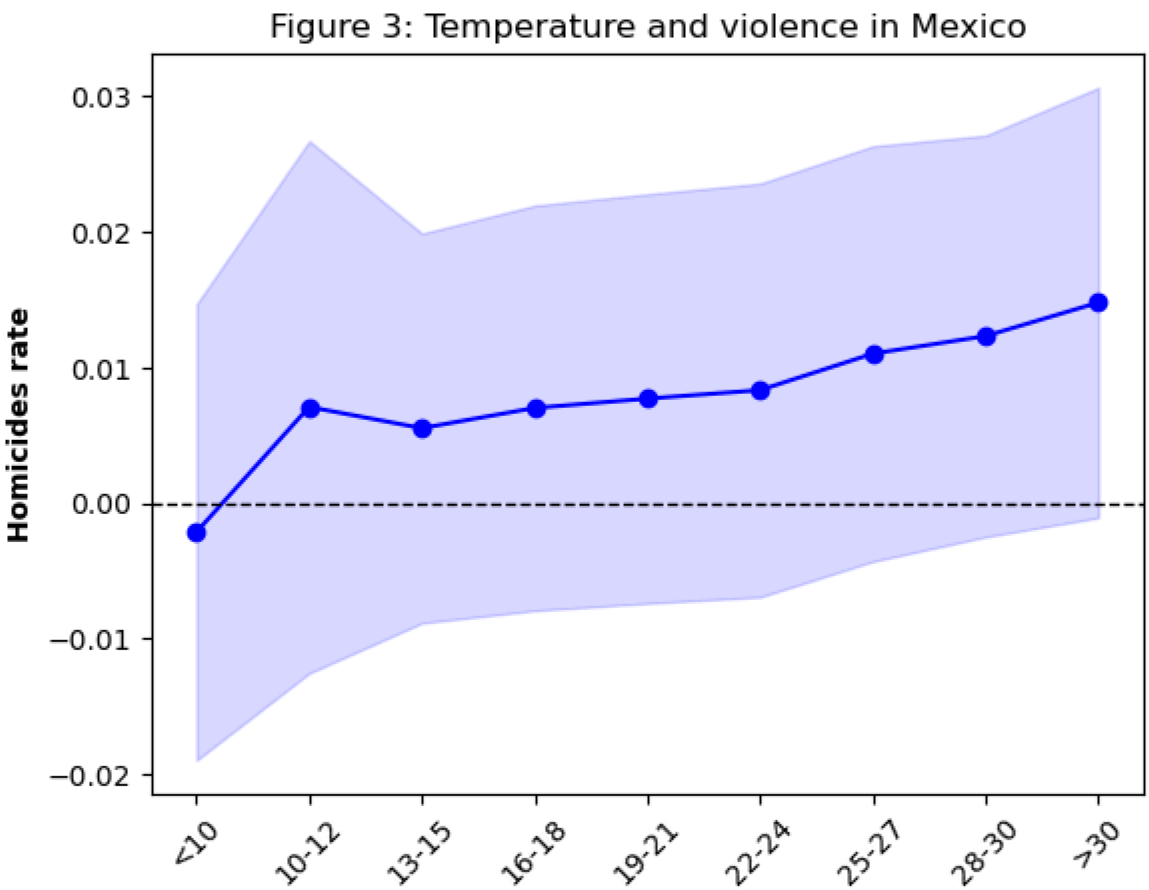


Figure 5: Temperature and suicides

In [11]:
```
# More Table Data cleaning: Filter for years and drop Nan rows with same temperatur
fig_5_df = muni[muni['year'] < 2007].dropna(subset = ['sui_rate', 'precip_pop'] + t

# Clustered around state, and have fixed effects for (muni_code, year, month)
fig_5_df = fig_5_df.set_index(['muni_code', 'date'])
fig_5_df = fig_5_df[['state', 'sui_rate', 'precip_pop', 'popw'] + t_bins]
fig_5_df.head()
```

Out[11]:

		state	sui_rate	precip_pop	popw	temp_bin_pop_0_9	te
muni_code	date						
1001	2001-04-01	Aguascalientes	0.150756	0.113894	655030.9375	0.0	
	1998-04-01	Aguascalientes	0.160717	0.000000	655030.9375	0.0	
	2001-11-01	Aguascalientes	0.446021	0.038703	655030.9375	0.0	
	2000-02-01	Aguascalientes	0.620399	0.002461	655030.9375	0.0	
	2001-07-01	Aguascalientes	0.449569	1.110245	655030.9375	0.0	

In [12]:

```
# Regressions with Time and Municipality Fixed Effects that are weighted population
fig5_mod = PanelOLS.from_formula('sui_rate ~ ' + ' + '.join(t_bins) + '+ precip_pop
fig5_results = fig5_mod.fit(cov_type='clustered', clusters = fig_5_df['state'])
fig5_coef = fig5_results.params
# 95% confidence level
fig5_ci = fig5_results.conf_int(level =  0.95)

fig5_result_df = pd.DataFrame({'coefficient': fig5_coef, 'CI_lower': fig5_ci['lower
fig5_result_df
```

Out[12]:

	coefficient	CI_lower	CI_upper
temp_bin_pop_0_9	-0.001114	-0.005392	0.003164
temp_bin_pop_10_12	0.001647	-0.001652	0.004946
temp_bin_pop_13_15	0.001935	-0.001005	0.004874
temp_bin_pop_16_18	0.002260	-0.000739	0.005258
temp_bin_pop_19_21	0.003365	0.000266	0.006464
temp_bin_pop_22_24	0.003625	0.000666	0.006584
temp_bin_pop_25_27	0.004514	0.001389	0.007640
temp_bin_pop_28_30	0.005622	0.002486	0.008759
temp_bin_pop_31_Inf	0.005633	0.001465	0.009801

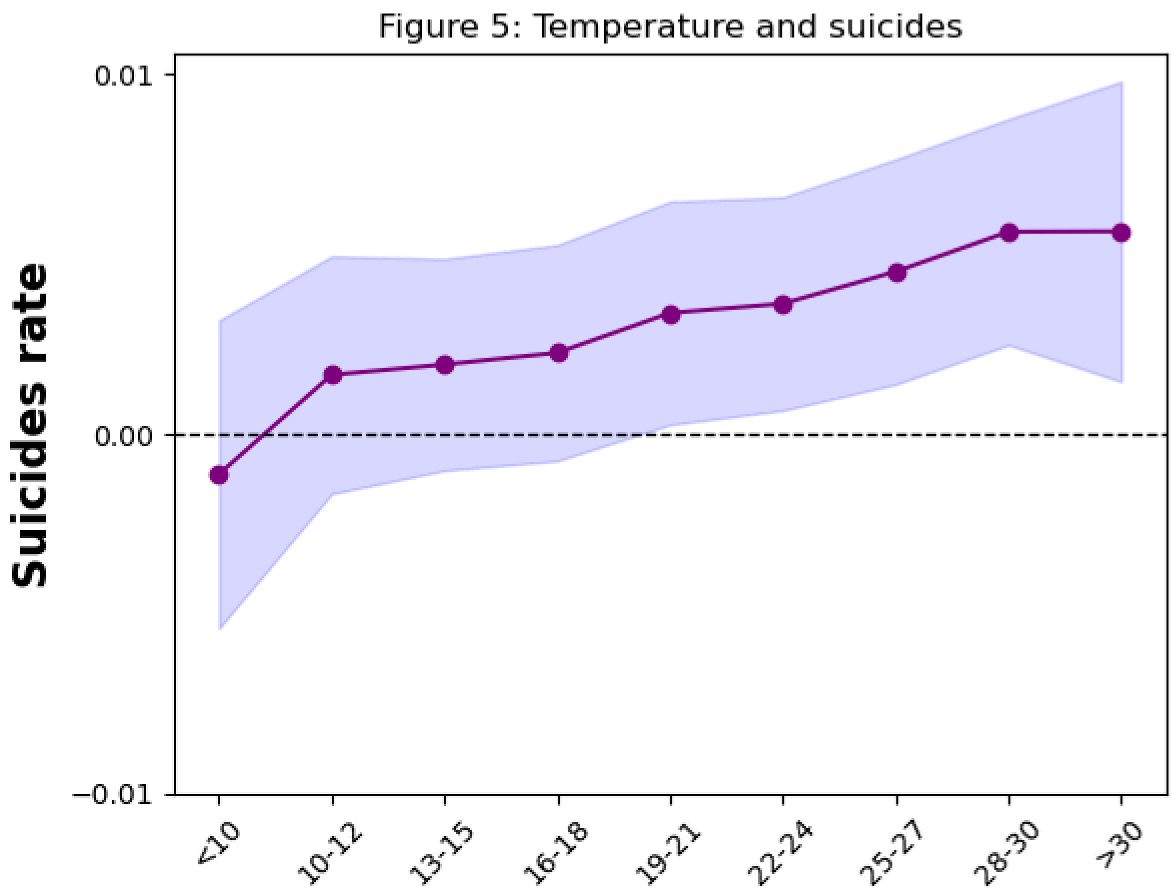
In [13]:

```
# Create the plot
fig, ax = plt.subplots()

# Plot ranges and points
ax.plot(fig5_result_df.index, fig5_result_df['coefficient'], '-o', color='purple')
ax.fill_between(fig5_result_df.index, fig5_result_df['CI_lower'], fig5_result_df['C
```

```
# Styling
ax.set_title('Figure 5: Temperature and suicides')
ax.axhline(0, color='black', linewidth=1, linestyle='dashed')
ax.set_xticks(fig5_result_df.index)
ax.set_xticklabels(["<10", "10-12", "13-15", "16-18", "19-21", "22-24", "25-27", "2
ax.set_ylabel('Suicides rate', fontweight='bold', fontsize = 16)
ax.set_xlabel('')
ax.set_yticks([-0.01, 0.00, 0.01])

# Show the plot
plt.show()
```



The data is different from what is being replicated (Width of 2 for temp_pop bins in the data vs Width of 3 bins shown in the paper)

Figure 6 : Temporal distribution of estimates

Homicides Rate

In [14]:
```
fig6_df = muni.copy()
for i in range(1, 7):
    fig6_df[f'L{i}_temp_pop'] = muni['temp_pop'].shift(-i)
    fig6_df[f'F{i}_temp_pop'] = muni['temp_pop'].shift(i)
    fig6_df[f'L{i}_precip_pop'] = muni['precip_pop'].shift(-i)
    fig6_df[f'F{i}_precip_pop'] = muni['precip_pop'].shift(i)
```

```
fig6_df
```

Out[14]:

	muni_code	state	municipality	year	month	temp_pop	precip_pop
0	1001	Aguascalientes	Aguascalientes	2001	4	23.817400	0.113894
1	1001	Aguascalientes	Aguascalientes	1998	4	22.207600	0.000000
2	1001	Aguascalientes	Aguascalientes	2001	11	16.335100	0.038703
3	1001	Aguascalientes	Aguascalientes	2000	2	17.702499	0.002461
4	1001	Aguascalientes	Aguascalientes	2001	7	21.856300	1.110245
...	...	...	...	...	...	...	...
618907	32058	Zacatecas	Santa María de la Paz	1994	5	26.773199	0.082976
618908	32058	Zacatecas	Santa María de la Paz	1994	10	23.050100	0.558740
618909	32058	Zacatecas	Santa María de la Paz	1997	11	19.530899	0.683117
618910	32058	Zacatecas	Santa María de la Paz	1994	11	20.543400	0.122095
618911	32058	Zacatecas	Santa María de la Paz	1996	2	19.468500	0.262075

618912 rows × 83 columns

In [15]:

```
temp_lags = ['L1_temp_pop', 'L2_temp_pop', 'L3_temp_pop', 'L4_temp_pop', 'L5_temp_p
temp_leads = ['F1_temp_pop', 'F2_temp_pop', 'F3_temp_pop', 'F4_temp_pop', 'F5_temp_
precip_lags = ['L1_precip_pop', 'L2_precip_pop', 'L3_precip_pop', 'L4_precip_pop',
precip_leads = ['F1_precip_pop', 'F2_precip_pop', 'F3_precip_pop', 'F4_precip_pop',

fig_6_hom_df = fig6_df[muni['year'] <= 2006].dropna(subset = ['hom_rate', 'temp_pop

# Clustered around state, and have fixed effects for (muni_code, year, month)
fig_6_hom_df = fig_6_hom_df.set_index(['muni_code', 'date'])
fig_6_hom_df = fig_6_hom_df[['state', 'hom_rate', 'temp_pop', 'precip_pop', 'popw',
fig_6_hom_df
```

Out[15]:

muni_code	date	state	hom_rate	temp_pop	precip_pop	popw	year	m
1001	1995-05-01	Aguascalientes	0.511190	24.609200	0.276077	655030.9375	1995	
	2004-03-01	Aguascalientes	0.000000	18.081400	0.224986	655030.9375	2004	
	2001-02-01	Aguascalientes	0.000000	17.777901	0.003232	655030.9375	2001	
	2004-11-01	Aguascalientes	0.138814	15.954700	0.044783	655030.9375	2004	
	2000-12-01	Aguascalientes	0.000000	14.781200	0.066093	655030.9375	2000	
...	...	...	...	...	...	...	...	
32058	2006-11-01	Zacatecas	0.000000	18.696400	0.192440	2711.0000	2006	
	2005-12-01	Zacatecas	0.000000	17.435801	0.000000	2711.0000	2005	
	2005-03-01	Zacatecas	38.355366	19.163700	0.044328	2711.0000	2005	
	2006-03-01	Zacatecas	0.000000	21.108900	0.002020	2711.0000	2006	
	2005-06-01	Zacatecas	0.000000	26.785601	0.608296	2711.0000	2005	

493862 rows × 31 columns

In [16]:

```
mod_fig_6_hom = PanelOLS.from_formula('hom_rate ~ temp_pop + ' + ' + '.join(temp_la
fig6_results = mod_fig_6_hom.fit(cov_type='clustered', clusters = fig_6_hom_df['sta

fig6_coef = fig6_results.params
fig6_ci = fig6_results.conf_int() # 95% confidence level

result_df_fig_6_hom = pd.DataFrame({'coefficient': fig6_coef, 'CI_lower': fig6_ci['

# Add the estimate together for the lags and leads for the plots
result_df_fig_6_hom = result_df_fig_6_hom.iloc[:-13, :].reindex(temp_lags[::-1] + [
result_df_fig_6_hom
```

Out[16]:

	coefficient	CI_lower	CI_upper
L6_temp_pop	-0.000799	-0.003210	0.001613
L5_temp_pop	-0.000194	-0.002995	0.002607
L4_temp_pop	-0.003508	-0.005920	-0.001097
L3_temp_pop	-0.001269	-0.003478	0.000939
L2_temp_pop	-0.000978	-0.004269	0.002313
L1_temp_pop	-0.002053	-0.004532	0.000425
temp_pop	0.015693	0.008152	0.023233
F1_temp_pop	-0.000493	-0.002661	0.001675
F2_temp_pop	0.001945	-0.000697	0.004587
F3_temp_pop	0.002411	-0.000944	0.005766
F4_temp_pop	0.002045	0.000166	0.003925
F5_temp_pop	-0.000621	-0.003371	0.002130
F6_temp_pop	0.001212	-0.001288	0.003713

Suicides Rate

In [17]:
```
fig_6_sui_df = fig6_df[fig6_df['year'] <= 2006].dropna(subset = ['sui_rate', 'temp_

fig_6_sui_df = fig_6_sui_df.set_index(['muni_code', 'date'])
fig_6_sui_df = fig_6_sui_df[['state', 'sui_rate', 'temp_pop', 'precip_pop', 'popw',
fig_6_sui_df
```

Out[17]:

		state	sui_rate	temp_pop	precip_pop	popw	year	mo
muni_code	**date**							
1001	**1995-05-01**	Aguascalientes	0.340793	24.609200	0.276077	655030.9375	1995	
	2004-03-01	Aguascalientes	0.422671	18.081400	0.224986	655030.9375	2004	
	2001-02-01	Aguascalientes	0.151361	17.777901	0.003232	655030.9375	2001	
	2004-11-01	Aguascalientes	0.694070	15.954700	0.044783	655030.9375	2004	
	2000-12-01	Aguascalientes	0.151972	14.781200	0.066093	655030.9375	2000	
...	**...**	...	...	...	...	...	...	
32058	**2006-11-01**	Zacatecas	0.000000	18.696400	0.192440	2711.0000	2006	
	2005-12-01	Zacatecas	0.000000	17.435801	0.000000	2711.0000	2005	
	2005-03-01	Zacatecas	0.000000	19.163700	0.044328	2711.0000	2005	
	2006-03-01	Zacatecas	0.000000	21.108900	0.002020	2711.0000	2006	
	2005-06-01	Zacatecas	0.000000	26.785601	0.608296	2711.0000	2005	

493862 rows × 31 columns

In [18]:

```
mod_fig_6_sui = PanelOLS.from_formula('sui_rate ~ temp_pop + ' + ' + '.join(temp_la
results_sui = mod_fig_6_sui.fit(cov_type='clustered', clusters = fig_6_sui_df['stat

coefficients_sui = results_sui.params
confidence_interval_sui = results_sui.conf_int() # 95% confidence level

result_df_fig_6_sui = pd.DataFrame({'coefficient': coefficients_sui, 'CI_lower': co

# Add the estimate together for the lags and leads for the plots
result_df_fig_6_sui = result_df_fig_6_sui.iloc[:-13, :].reindex(temp_lags[::-1] + [
result_df_fig_6_sui
```

Out[18]:

	coefficient	CI_lower	CI_upper
L6_temp_pop	-0.000551	-0.001558	0.000455
L5_temp_pop	-0.000051	-0.000862	0.000761
L4_temp_pop	-0.000393	-0.000997	0.000211
L3_temp_pop	-0.000195	-0.000809	0.000419
L2_temp_pop	-0.000118	-0.001220	0.000984
L1_temp_pop	-0.000003	-0.001066	0.001060
temp_pop	0.007186	0.005241	0.009130
F1_temp_pop	0.000063	-0.000788	0.000913
F2_temp_pop	-0.000950	-0.002253	0.000353
F3_temp_pop	0.000137	-0.001282	0.001557
F4_temp_pop	0.000511	-0.000198	0.001221
F5_temp_pop	0.000017	-0.000948	0.000982
F6_temp_pop	-0.000606	-0.001656	0.000445

In [19]:

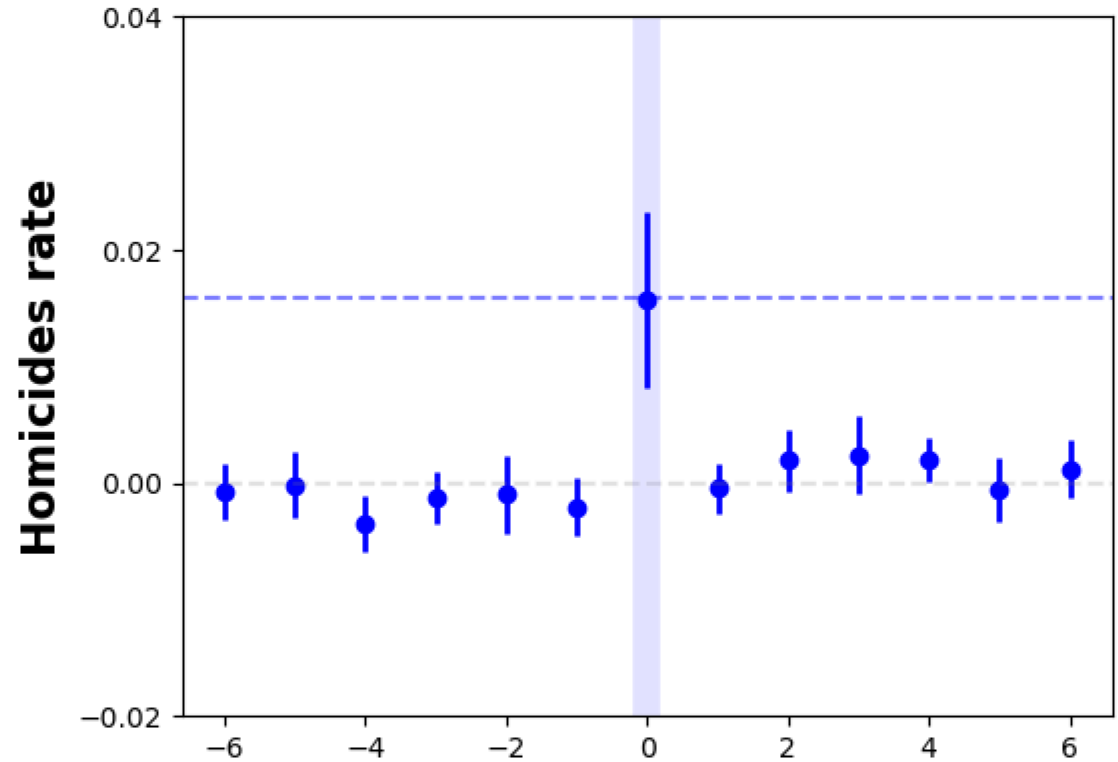
```
x = np.arange(-6, 7)

fig, (ax1, ax2) = plt.subplots(2, 1, figsize=(6, 10))

ax1.scatter(x, result_df_fig_6_hom['coefficient'], color='blue')
ax1.vlines(x, result_df_fig_6_hom['CI_lower'], result_df_fig_6_hom['CI_upper'], col
ax1.axhline(0, linestyle='--', color='gray', alpha = 0.2)
ax1.axhline(0.0158, linestyle='--', color='blue', alpha = 0.5)
ax1.axvline(0, linewidth = 10, color='blue', alpha = 0.1)
ax1.set_xlabel('')
ax1.set_ylabel('Homicides rate', fontweight='bold', fontsize= 16)
ax1.set_yticks([-.02, 0, .02, .04])

ax2.scatter(x, result_df_fig_6_sui['coefficient'], color='purple')
ax2.vlines(x, result_df_fig_6_sui['CI_lower'], result_df_fig_6_sui['CI_upper'], col
ax2.axhline(0, linestyle='--', color='gray', alpha = 0.2)
ax2.axhline(0.0070, linestyle='--', color='purple', alpha = 0.5)
ax2.axvline(0, linewidth = 10, color='purple', alpha = 0.1)
ax2.set_xlabel('')
ax2.set_ylabel('Suicides rate', fontweight='bold', fontsize= 16)
ax2.set_yticks([-.005, 0, .005, .01, .015])

plt.show()
```



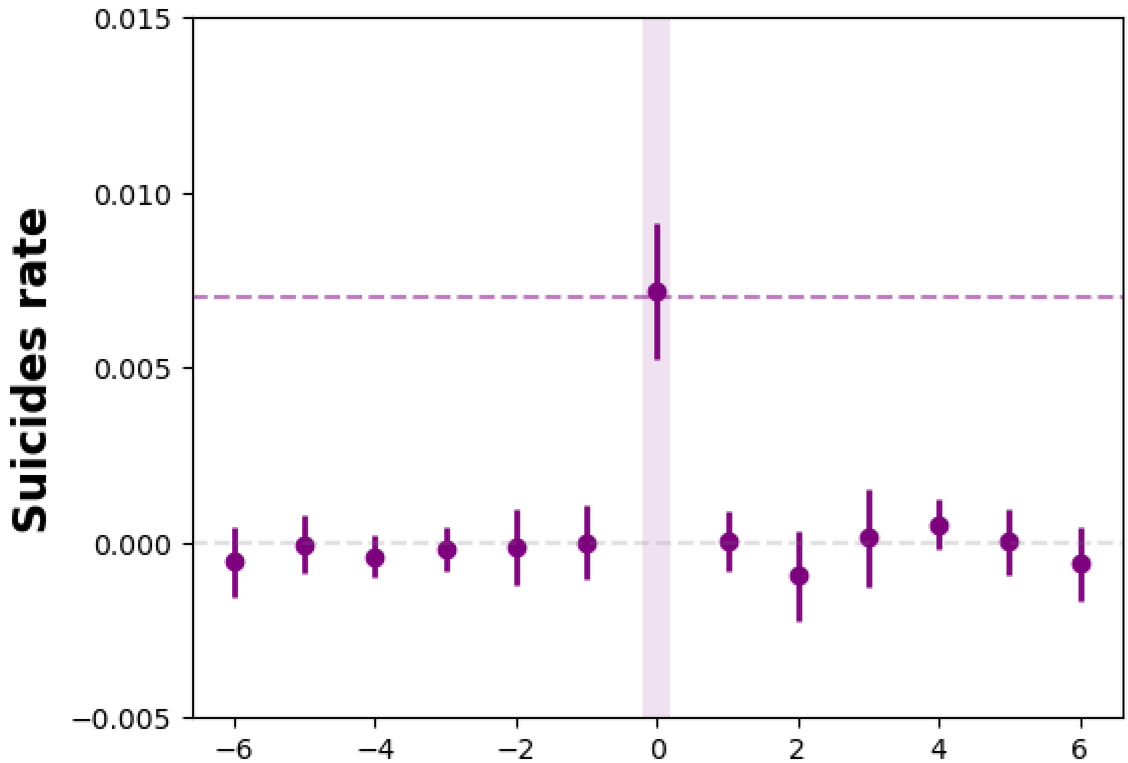


Table 1: Descriptive Statistics

In [20]:
```
vars = ['hom_rate', 'sui_rate', 'population', 'temp_pop', 'precip_pop']
tbl1_1 = muni[muni['year'] < 2007][vars].describe().loc[['mean', 'std', 'count']].a
tbl1_1
```

Out[20]:

	hom_rate	sui_rate	population	temp_pop	precip_pop
mean	0.98	0.21	39057.12	20.10	0.93
std	5.23	1.92	116901.42	5.00	1.12
count	494724.00	494724.00	494724.00	500208.00	500208.00

In [21]:
```
vars = ['hom_rate', 'sui_rate', 'population', 'temp_pop', 'precip_pop']
tbl1_2 = muni[muni['year'] >= 2007][vars].describe().loc[['mean', 'std', 'count']].
tbl1_2
```

Out[21]:

	hom_rate	sui_rate	population	temp_pop	precip_pop
mean	0.83	0.26	44584.62	20.05	0.81
std	4.13	2.21	130760.44	5.09	1.07
count	117650.00	117650.00	117650.00	117696.00	117696.00

In [22]:
```
tbl1_df_filtered = muni[muni['year'] < 2007]
tbl1_df = tbl1_df_filtered.groupby('muni_code')['hom_rate'].apply(list)
x = f_oneway(tbl1_df[1001],tbl1_df[1002],tbl1_df[1003],tbl1_df[1004],tbl1_df[1005],
print(x)
```

```
F_onewayResult(statistic=nan, pvalue=nan)
```

Cannot run f_oneway for data with more than 9 groups

Table 2: Temperature and Violence in Mexico

In [23]:
```
tbl2_df = muni[muni['year'] < 2007].dropna().set_index(['muni_code', 'date'])[['tem
tbl2_df.head()
```

Out[23]:

muni_code	date	temp_pop	state	hom_rate	precip_pop	popw
1001	**2001-04-01**	23.817400	Aguascalientes	0.301511	0.113894	655030.9375
	1998-04-01	22.207600	Aguascalientes	0.642868	0.000000	655030.9375
	2001-11-01	16.335100	Aguascalientes	0.297347	0.038703	655030.9375
	2000-02-01	17.702499	Aguascalientes	0.000000	0.002461	655030.9375
	2001-07-01	21.856300	Aguascalientes	0.149856	1.110245	655030.9375

In [24]:
```
# Column 5
col5_model = PanelOLS.from_formula('hom_rate ~ temp_pop + precip_pop + EntityEffect
col5_results = col5_model.fit(cov_type='clustered', clusters = tbl2_df['state'])

print(col5_results)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0005
Estimator:                   PanelOLS   R-squared (Between):              0.1645
No. Observations:              119136   R-squared (Within):              -0.0001
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.1004
Time:                        15:10:14   Log-likelihood                -2.159e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      27.802
Entities:                         584   P-value                           0.0000
Avg Obs:                       204.00   Distribution:                F(2,118347)
Min Obs:                       204.00
Max Obs:                       204.00   F-statistic (robust):             10.238
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(2,118347)
Avg Obs:                       584.00
Min Obs:                       584.00
Max Obs:                       584.00

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0164     0.0040     4.0847     0.0000      0.0085      0.0243
precip_pop    -0.0050     0.0089    -0.5637     0.5729     -0.0225      0.0125
==============================================================================

F-test for Poolability: 155.29
P-value: 0.0000
Distribution: F(786,118347)

Included effects: Entity, Time
```

Table 3: Temperature and Economically Motivated Crimes

In [25]:
```python
tbl3_df = state[state['year'] >= 2007]
tbl3_df['month'] = pd.Categorical(tbl3_df['month'])
tbl3_df['year'] = pd.Categorical(tbl3_df['year'])
tbl3_df
```

```
C:\Users\blahb\AppData\Local\Temp\ipykernel_16636\686763787.py:2: SettingWithCopyWar
ning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/u
ser_guide/indexing.html#returning-a-view-versus-a-copy
  tbl3_df['month'] = pd.Categorical(tbl3_df['month'])
C:\Users\blahb\AppData\Local\Temp\ipykernel_16636\686763787.py:3: SettingWithCopyWar
ning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/u
ser_guide/indexing.html#returning-a-view-versus-a-copy
  tbl3_df['year'] = pd.Categorical(tbl3_df['year'])
```

Out[25]:

	id	year	state	month	temperature	homsegob_rate	kid_rate	ext_rat
204	1804.0	2007	Aguascalientes	1	13.600000	0.178575	0.000000	0.62501
205	1804.0	2007	Aguascalientes	2	14.566670	0.535724	0.000000	0.17857
206	1804.0	2007	Aguascalientes	3	18.100000	0.089287	0.178575	0.44643
207	1804.0	2007	Aguascalientes	4	19.299999	0.178575	0.267862	0.17857
208	1804.0	2007	Aguascalientes	5	22.066669	0.357149	0.267862	0.26786
...	...	...	...	...	...	...	...	
8059	1835.0	2010	Zacatecas	8	21.160000	0.402504	0.000000	0.06708
8060	1835.0	2010	Zacatecas	9	19.936001	0.603756	0.000000	0.13416
8061	1835.0	2010	Zacatecas	10	17.436001	0.805008	0.067084	0.06708
8062	1835.0	2010	Zacatecas	11	14.444000	0.872092	0.268336	0.06708
8063	1835.0	2010	Zacatecas	12	12.864000	0.939176	0.067084	0.20125

1536 rows × 12 columns

In [26]:

```
model_a = sm.OLS.from_formula('homsegob_rate ~ temperature + precipitation + C(mont
result_a = model_a.fit()

model_b = sm.OLS.from_formula('car_rate ~ temperature + precipitation + C(month) +
```

```
result_b = model_b.fit()

model_c = sm.OLS.from_formula('ext_rate ~ temperature + precipitation + C(month) +
result_c = model_c.fit()

model_d = sm.OLS.from_formula('kid_rate ~ temperature + precipitation + C(month) +
result_d = model_d.fit()
```

In [27]:
```
print(result_a.summary())
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:          homsegob_rate   R-squared:                       0.714
Model:                            OLS   Adj. R-squared:                  0.705
Method:                 Least Squares   F-statistic:                     78.99
Date:                Fri, 26 Apr 2024   Prob (F-statistic):               0.00
Time:                        15:10:14   Log-Likelihood:                -1876.4
No. Observations:                1535   AIC:                             3849.
Df Residuals:                    1487   BIC:                             4105.
Df Model:                          47
Covariance Type:            nonrobust
====================================================================================
                       coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------
Intercept           -0.8281      0.224     -3.705      0.000      -1.267      -0.390
C(month)[T.2]       -0.1004      0.107     -0.938      0.348      -0.310       0.110
C(month)[T.3]       -0.0558      0.113     -0.493      0.622      -0.277       0.166
C(month)[T.4]       -0.1996      0.127     -1.569      0.117      -0.449       0.050
C(month)[T.5]       -0.1452      0.146     -0.996      0.320      -0.431       0.141
C(month)[T.6]       -0.1176      0.164     -0.719      0.472      -0.438       0.203
C(month)[T.7]       -0.1635      0.140     -1.169      0.243      -0.438       0.111
C(month)[T.8]       -0.0748      0.173     -0.433      0.665      -0.414       0.264
C(month)[T.9]       -0.0854      0.169     -0.505      0.613      -0.417       0.246
C(month)[T.10]       0.0217      0.130      0.168      0.867      -0.233       0.276
C(month)[T.11]       0.0824      0.109      0.755      0.450      -0.132       0.297
C(month)[T.12]       0.2036      0.105      1.935      0.053      -0.003       0.410
C(year)[T.2008]      0.2634      0.060      4.368      0.000       0.145       0.382
C(year)[T.2009]      0.5032      0.060      8.335      0.000       0.385       0.622
C(year)[T.2010]      0.9733      0.060     16.108      0.000       0.855       1.092
C(id)[T.1805.0]      1.6654      0.172      9.676      0.000       1.328       2.003
C(id)[T.1806.0]     -0.2058      0.178     -1.159      0.247      -0.554       0.143
C(id)[T.1807.0]     -0.3477      0.200     -1.738      0.082      -0.740       0.045
C(id)[T.1808.0]      0.0595      0.194      0.307      0.759      -0.320       0.439
C(id)[T.1809.0]      5.6270      0.172     32.713      0.000       5.290       5.964
C(id)[T.1810.0]      0.1451      0.173      0.838      0.402      -0.195       0.485
C(id)[T.1811.0]     -0.1424      0.197     -0.722      0.470      -0.529       0.244
C(id)[T.1812.0]      0.5529      0.185      2.991      0.003       0.190       0.916
C(id)[T.1813.0]      2.9995      0.171     17.540      0.000       2.664       3.335
C(id)[T.1814.0]      0.0640      0.170      0.376      0.707      -0.270       0.398
C(id)[T.1815.0]      2.1814      0.187     11.685      0.000       1.815       2.548
C(id)[T.1816.0]      0.0471      0.171      0.276      0.783      -0.288       0.383
C(id)[T.1817.0]      0.0845      0.175      0.484      0.629      -0.258       0.427
C(id)[T.1818.0]      0.3671      0.173      2.118      0.034       0.027       0.707
C(id)[T.1819.0]      0.5848      0.177      3.313      0.001       0.239       0.931
C(id)[T.1820.0]      0.6981      0.182      3.842      0.000       0.342       1.054
C(id)[T.1821.0]      0.8811      0.179      4.915      0.000       0.529       1.233
C(id)[T.1822.0]      0.1845      0.173      1.066      0.286      -0.155       0.524
C(id)[T.1823.0]      1.1031      0.180      6.112      0.000       0.749       1.457
C(id)[T.1824.0]      0.1025      0.174      0.589      0.556      -0.239       0.444
C(id)[T.1825.0]     -0.0282      0.173     -0.163      0.870      -0.367       0.310
C(id)[T.1826.0]      0.4052      0.210      1.929      0.054      -0.007       0.817
C(id)[T.1827.0]      0.2336      0.172      1.358      0.175      -0.104       0.571
C(id)[T.1828.0]      3.3408      0.188     17.806      0.000       2.973       3.709
C(id)[T.1829.0]      0.8226      0.177      4.644      0.000       0.475       1.170
C(id)[T.1830.0]     -0.2732      0.218     -1.254      0.210      -0.701       0.154
C(id)[T.1831.0]      0.3414      0.182      1.880      0.060      -0.015       0.698
```

```
C(id)[T.1832.0]    -0.1043      0.172     -0.606      0.544      -0.442       0.233
C(id)[T.1833.0]    -0.2099      0.198     -1.061      0.289      -0.598       0.178
C(id)[T.1834.0]    -0.7318      0.209     -3.497      0.000      -1.142      -0.321
C(id)[T.1835.0]     0.0367      0.171      0.215      0.830      -0.298       0.371
temperature         0.0500      0.013      3.946      0.000       0.025       0.075
precipitation      -0.2852      0.503     -0.567      0.571      -1.273       0.702
==============================================================================
Omnibus:                      333.910   Durbin-Watson:                   0.563
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             6278.150
Skew:                           0.496   Prob(JB):                         0.00
Kurtosis:                      12.858   Cond. No.                         759.
==============================================================================

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

In [28]:

```
print(result_b.summary())
```

```
                            OLS Regression Results                            
==============================================================================
Dep. Variable:               car_rate   R-squared:                       0.886
Model:                            OLS   Adj. R-squared:                  0.882
Method:                 Least Squares   F-statistic:                     245.2
Date:                Fri, 26 Apr 2024   Prob (F-statistic):               0.00
Time:                        15:10:14   Log-Likelihood:                -4821.7
No. Observations:                1535   AIC:                             9739.
Df Residuals:                    1487   BIC:                             9996.
Df Model:                          47                                         
Covariance Type:            nonrobust                                         
=====================================================================================
                        coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------------
Intercept             9.1116      1.523      5.984      0.000       6.125      12.099
C(month)[T.2]        -0.7535      0.729     -1.033      0.302      -2.184       0.677
C(month)[T.3]        -0.0992      0.770     -0.129      0.897      -1.610       1.411
C(month)[T.4]        -1.0083      0.866     -1.164      0.245      -2.707       0.691
C(month)[T.5]        -0.3103      0.994     -0.312      0.755      -2.260       1.639
C(month)[T.6]        -0.3848      1.114     -0.345      0.730      -2.570       1.800
C(month)[T.7]         0.0202      0.953      0.021      0.983      -1.848       1.889
C(month)[T.8]         0.3539      1.177      0.301      0.764      -1.955       2.663
C(month)[T.9]         0.3914      1.151      0.340      0.734      -1.867       2.650
C(month)[T.10]        1.0691      0.884      1.209      0.227      -0.665       2.803
C(month)[T.11]        0.7541      0.744      1.014      0.311      -0.704       2.213
C(month)[T.12]        1.1761      0.717      1.641      0.101      -0.230       2.582
C(year)[T.2008]       2.2378      0.411      5.449      0.000       1.432       3.043
C(year)[T.2009]       2.2570      0.411      5.488      0.000       1.450       3.064
C(year)[T.2010]       4.7792      0.412     11.611      0.000       3.972       5.587
C(id)[T.1805.0]      65.7676      1.173     56.091      0.000      63.468      68.068
C(id)[T.1806.0]       1.5914      1.210      1.315      0.189      -0.782       3.965
C(id)[T.1807.0]     -12.4254      1.363     -9.116      0.000     -15.099      -9.752
C(id)[T.1808.0]     -10.4977      1.319     -7.959      0.000     -13.085      -7.910
C(id)[T.1809.0]      43.6889      1.172     37.283      0.000      41.390      45.987
C(id)[T.1810.0]      -6.6415      1.180     -5.627      0.000      -8.956      -4.326
C(id)[T.1811.0]      -7.2718      1.343     -5.413      0.000      -9.907      -4.637
C(id)[T.1812.0]      12.0962      1.259      9.605      0.000       9.626      14.567
C(id)[T.1813.0]       3.0583      1.165      2.625      0.009       0.773       5.344
C(id)[T.1814.0]      -6.4172      1.161     -5.527      0.000      -8.695      -4.140
C(id)[T.1815.0]      -5.4527      1.272     -4.287      0.000      -7.947      -2.958
C(id)[T.1816.0]      -4.8264      1.165     -4.143      0.000      -7.112      -2.541
C(id)[T.1817.0]      -4.1277      1.190     -3.469      0.001      -6.462      -1.793
C(id)[T.1818.0]       8.5017      1.181      7.200      0.000       6.186      10.818
C(id)[T.1819.0]      -3.8846      1.203     -3.230      0.001      -6.243      -1.526
C(id)[T.1820.0]       2.0046      1.238      1.620      0.106      -0.423       4.432
C(id)[T.1821.0]      -6.4401      1.221     -5.273      0.000      -8.836      -4.045
C(id)[T.1822.0]       9.2310      1.179      7.833      0.000       6.919      11.543
C(id)[T.1823.0]      -9.9065      1.230     -8.057      0.000     -12.318      -7.495
C(id)[T.1824.0]      -7.3084      1.185     -6.169      0.000      -9.632      -4.984
C(id)[T.1825.0]      -4.0836      1.175     -3.475      0.001      -6.389      -1.778
C(id)[T.1826.0]      -5.9988      1.431     -4.192      0.000      -8.806      -3.192
C(id)[T.1827.0]      -6.5833      1.172     -5.616      0.000      -8.883      -4.284
C(id)[T.1828.0]       6.5008      1.278      5.086      0.000       3.994       9.008
C(id)[T.1829.0]       1.8938      1.207      1.570      0.117      -0.473       4.261
C(id)[T.1830.0]      -7.4100      1.484     -4.992      0.000     -10.322      -4.498
C(id)[T.1831.0]       5.1062      1.237      4.127      0.000       2.679       7.533
```

```
C(id)[T.1832.0]    -8.4107      1.172     -7.178      0.000     -10.709      -6.112
C(id)[T.1833.0]    -9.3248      1.348     -6.917      0.000     -11.969      -6.681
C(id)[T.1834.0]   -11.7955      1.426     -8.274      0.000     -14.592      -8.999
C(id)[T.1835.0]    -3.9206      1.162     -3.374      0.001      -6.200      -1.641
temperature         0.0666      0.086      0.771      0.441      -0.103       0.236
precipitation      -0.3632      3.429     -0.106      0.916      -7.090       6.363
==============================================================================
Omnibus:                      330.297   Durbin-Watson:                   0.296
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            10136.823
Skew:                          -0.241   Prob(JB):                         0.00
Kurtosis:                      15.580   Cond. No.                         759.
==============================================================================

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

In [29]:

```
print(result_c.summary())
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:               ext_rate   R-squared:                       0.603
Model:                            OLS   Adj. R-squared:                  0.590
Method:                 Least Squares   F-statistic:                     47.96
Date:                Fri, 26 Apr 2024   Prob (F-statistic):         1.45e-260
Time:                        15:10:14   Log-Likelihood:                -605.89
No. Observations:                1535   AIC:                             1308.
Df Residuals:                    1487   BIC:                             1564.
Df Model:                          47
Covariance Type:            nonrobust
=====================================================================================
                        coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------------
Intercept             0.2571      0.098      2.632      0.009       0.066       0.449
C(month)[T.2]        -0.0075      0.047     -0.161      0.873      -0.099       0.084
C(month)[T.3]         0.0421      0.049      0.853      0.394      -0.055       0.139
C(month)[T.4]         0.0931      0.056      1.675      0.094      -0.016       0.202
C(month)[T.5]         0.0402      0.064      0.630      0.529      -0.085       0.165
C(month)[T.6]         0.0201      0.071      0.282      0.778      -0.120       0.160
C(month)[T.7]         0.0173      0.061      0.284      0.777      -0.103       0.137
C(month)[T.8]        -0.0195      0.076     -0.259      0.796      -0.168       0.129
C(month)[T.9]         0.0167      0.074      0.226      0.821      -0.128       0.162
C(month)[T.10]        0.0413      0.057      0.729      0.466      -0.070       0.153
C(month)[T.11]       -0.0125      0.048     -0.263      0.793      -0.106       0.081
C(month)[T.12]       -0.0270      0.046     -0.587      0.557      -0.117       0.063
C(year)[T.2008]       0.1273      0.026      4.832      0.000       0.076       0.179
C(year)[T.2009]       0.2683      0.026     10.167      0.000       0.217       0.320
C(year)[T.2010]       0.2273      0.026      8.608      0.000       0.175       0.279
C(id)[T.1805.0]       0.6657      0.075      8.850      0.000       0.518       0.813
C(id)[T.1806.0]       0.5423      0.078      6.987      0.000       0.390       0.695
C(id)[T.1807.0]      -0.3062      0.087     -3.502      0.000      -0.478      -0.135
C(id)[T.1808.0]      -0.1635      0.085     -1.933      0.053      -0.330       0.002
C(id)[T.1809.0]       0.2027      0.075      2.696      0.007       0.055       0.350
C(id)[T.1810.0]      -0.2041      0.076     -2.696      0.007      -0.353      -0.056
C(id)[T.1811.0]      -0.3009      0.086     -3.491      0.000      -0.470      -0.132
C(id)[T.1812.0]       0.4076      0.081      5.045      0.000       0.249       0.566
C(id)[T.1813.0]       0.3214      0.075      4.300      0.000       0.175       0.468
C(id)[T.1814.0]       0.0110      0.074      0.148      0.882      -0.135       0.157
C(id)[T.1815.0]      -0.1735      0.082     -2.127      0.034      -0.334      -0.013
C(id)[T.1816.0]       0.0359      0.075      0.480      0.631      -0.111       0.182
C(id)[T.1817.0]       0.4276      0.076      5.601      0.000       0.278       0.577
C(id)[T.1818.0]      -0.3581      0.076     -4.728      0.000      -0.507      -0.210
C(id)[T.1819.0]       0.1647      0.077      2.135      0.033       0.013       0.316
C(id)[T.1820.0]       1.7288      0.079     21.773      0.000       1.573       1.885
C(id)[T.1821.0]      -0.3188      0.078     -4.070      0.000      -0.473      -0.165
C(id)[T.1822.0]      -0.2427      0.076     -3.209      0.001      -0.391      -0.094
C(id)[T.1823.0]       0.5821      0.079      7.380      0.000       0.427       0.737
C(id)[T.1824.0]      -0.3403      0.076     -4.478      0.000      -0.489      -0.191
C(id)[T.1825.0]      -0.1864      0.075     -2.472      0.014      -0.334      -0.039
C(id)[T.1826.0]       0.3797      0.092      4.136      0.000       0.200       0.560
C(id)[T.1827.0]       0.5119      0.075      6.807      0.000       0.364       0.659
C(id)[T.1828.0]      -0.0879      0.082     -1.072      0.284      -0.249       0.073
C(id)[T.1829.0]      -0.1985      0.077     -2.565      0.010      -0.350      -0.047
C(id)[T.1830.0]       0.3176      0.095      3.336      0.001       0.131       0.504
C(id)[T.1831.0]      -0.0539      0.079     -0.678      0.498      -0.210       0.102
```

```
C(id)[T.1832.0]    -0.3515      0.075     -4.676      0.000      -0.499      -0.204
C(id)[T.1833.0]     0.0422      0.086      0.488      0.626      -0.127       0.212
C(id)[T.1834.0]    -0.2956      0.091     -3.231      0.001      -0.475      -0.116
C(id)[T.1835.0]    -0.2002      0.075     -2.686      0.007      -0.346      -0.054
temperature        -0.0055      0.006     -0.989      0.323      -0.016       0.005
precipitation       0.2203      0.220      1.001      0.317      -0.211       0.652
==============================================================================
Omnibus:                     1193.799   Durbin-Watson:                   0.950
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            56571.437
Skew:                           3.185   Prob(JB):                         0.00
Kurtosis:                      32.051   Cond. No.                         759.
==============================================================================

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

In [30]:
```
print(result_d.summary())
```

```
                            OLS Regression Results                            
==============================================================================
Dep. Variable:               kid_rate   R-squared:                       0.392
Model:                            OLS   Adj. R-squared:                  0.372
Method:                 Least Squares   F-statistic:                     20.34
Date:                Fri, 26 Apr 2024   Prob (F-statistic):          3.79e-127
Time:                        15:10:14   Log-Likelihood:                 1538.7
No. Observations:                1534   AIC:                            -2981.
Df Residuals:                    1486   BIC:                            -2725.
Df Model:                          47                                         
Covariance Type:            nonrobust                                         
=====================================================================================
                        coef    std err          t      P>|t|      [0.025      0.975]
-------------------------------------------------------------------------------------
Intercept             0.0809      0.024      3.349      0.001       0.034       0.128
C(month)[T.2]         0.0020      0.012      0.169      0.866      -0.021       0.025
C(month)[T.3]         0.0242      0.012      1.985      0.047       0.000       0.048
C(month)[T.4]         0.0080      0.014      0.584      0.559      -0.019       0.035
C(month)[T.5]         0.0127      0.016      0.805      0.421      -0.018       0.044
C(month)[T.6]        -0.0008      0.018     -0.046      0.963      -0.035       0.034
C(month)[T.7]         0.0199      0.015      1.318      0.188      -0.010       0.050
C(month)[T.8]        -0.0028      0.019     -0.151      0.880      -0.039       0.034
C(month)[T.9]        -0.0019      0.018     -0.105      0.916      -0.038       0.034
C(month)[T.10]     9.874e-05      0.014      0.007      0.994      -0.027       0.028
C(month)[T.11]        0.0160      0.012      1.357      0.175      -0.007       0.039
C(month)[T.12]        0.0057      0.011      0.505      0.614      -0.017       0.028
C(year)[T.2008]       0.0366      0.007      5.610      0.000       0.024       0.049
C(year)[T.2009]       0.0597      0.007      9.142      0.000       0.047       0.072
C(year)[T.2010]       0.0656      0.007     10.042      0.000       0.053       0.078
C(id)[T.1805.0]       0.0978      0.019      5.263      0.000       0.061       0.134
C(id)[T.1806.0]      -0.0834      0.019     -4.347      0.000      -0.121      -0.046
C(id)[T.1807.0]      -0.1006      0.022     -4.656      0.000      -0.143      -0.058
C(id)[T.1808.0]      -0.0908      0.021     -4.343      0.000      -0.132      -0.050
C(id)[T.1809.0]       0.1876      0.019     10.098      0.000       0.151       0.224
C(id)[T.1810.0]      -0.0285      0.019     -1.523      0.128      -0.065       0.008
C(id)[T.1811.0]      -0.0821      0.021     -3.854      0.000      -0.124      -0.040
C(id)[T.1812.0]      -0.0291      0.020     -1.459      0.145      -0.068       0.010
C(id)[T.1813.0]       0.0397      0.018      2.147      0.032       0.003       0.076
C(id)[T.1814.0]      -0.0581      0.018     -3.156      0.002      -0.094      -0.022
C(id)[T.1815.0]      -0.0168      0.020     -0.831      0.406      -0.056       0.023
C(id)[T.1816.0]      -0.0638      0.018     -3.455      0.001      -0.100      -0.028
C(id)[T.1817.0]      -0.0976      0.019     -5.172      0.000      -0.135      -0.061
C(id)[T.1818.0]      -0.0540      0.019     -2.883      0.004      -0.091      -0.017
C(id)[T.1819.0]       0.0552      0.019      2.897      0.004       0.018       0.093
C(id)[T.1820.0]      -0.0375      0.020     -1.911      0.056      -0.076       0.001
C(id)[T.1821.0]      -0.0973      0.019     -4.998      0.000      -0.135      -0.059
C(id)[T.1822.0]      -0.0940      0.019     -5.030      0.000      -0.131      -0.057
C(id)[T.1823.0]      -0.0665      0.019     -3.411      0.001      -0.105      -0.028
C(id)[T.1824.0]      -0.0951      0.019     -5.065      0.000      -0.132      -0.058
C(id)[T.1825.0]      -0.1045      0.019     -5.607      0.000      -0.141      -0.068
C(id)[T.1826.0]      -0.0394      0.023     -1.736      0.083      -0.084       0.005
C(id)[T.1827.0]      -0.0884      0.019     -4.757      0.000      -0.125      -0.052
C(id)[T.1828.0]      -0.0679      0.020     -3.349      0.001      -0.108      -0.028
C(id)[T.1829.0]      -0.0965      0.019     -5.043      0.000      -0.134      -0.059
C(id)[T.1830.0]      -0.0378      0.024     -1.605      0.109      -0.084       0.008
C(id)[T.1831.0]      -0.0256      0.020     -1.305      0.192      -0.064       0.013
```

```
C(id)[T.1832.0]    -0.1134      0.019     -6.103      0.000      -0.150      -0.077
C(id)[T.1833.0]    -0.1086      0.021     -5.083      0.000      -0.151      -0.067
C(id)[T.1834.0]    -0.1148      0.023     -5.078      0.000      -0.159      -0.070
C(id)[T.1835.0]    -0.0128      0.018     -0.695      0.487      -0.049       0.023
temperature        -0.0007      0.001     -0.475      0.635      -0.003       0.002
precipitation       0.0603      0.054      1.109      0.268      -0.046       0.167
==============================================================================
Omnibus:                      821.064   Durbin-Watson:                   1.165
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            10489.425
Skew:                           2.208   Prob(JB):                         0.00
Kurtosis:                      15.025   Cond. No.                         759.
==============================================================================

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

Table 4: Interaction with Economic Variables

In [31]:
```
#TABLE 4 COLUMN 5
table_4_inc_df = muni[muni['year'] <= 2006].dropna(subset = ['hom_rate', 'temp_pop'
table_4_inc_df = table_4_inc_df.set_index(['muni_code', 'date'])
table_4_inc_df = table_4_inc_df[['state', 'hom_rate', 'popw', 'temp_pop', 'precip_p
table_4_inc_df
```

Out[31]:

muni_code	date	state	hom_rate	popw	temp_pop	precip_pop	temp_inc
1001	2001-04-01	Aguascalientes	0.301511	655030.9375	23.817400	0.113894	2.53
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	0.000000	1.67
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	0.038703	-1.44
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	0.002461	-0.71
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	1.110245	1.49
...	...	...	...	...	...	...	
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	1.465000	1.39
	1993-03-01	Zacatecas	2.608534	124968.4375	13.100000	0.016000	-0.59
	2003-07-01	Zacatecas	0.000000	124968.4375	16.400000	1.907000	0.22
	2005-12-01	Zacatecas	1.503913	124968.4375	11.900000	0.031000	-0.89
	2004-11-01	Zacatecas	1.517868	124968.4375	13.100000	0.029000	-0.59

486132 rows × 6 columns

In [32]:

```
mod_inc = PanelOLS.from_formula('hom_rate ~ temp_pop + precip_pop + temp_income + T

result_inc = mod_inc.fit(cov_type='clustered', clusters = table_4_inc_df['state'])

print(result_inc.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1482
No. Observations:              486132   R-squared (Within):              -0.0001
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0669
Time:                        15:11:42   Log-likelihood                -1.034e+06
Cov. Estimator:             Clustered
                                        F-statistic:                      33.295
Entities:                        2383   P-value                           0.0000
Avg Obs:                       204.00   Distribution:               F(3,483543)
Min Obs:                       204.00
Max Obs:                       204.00   F-statistic (robust):             7.6896
                                        P-value                           0.0000
Time periods:                     204   Distribution:               F(3,483543)
Avg Obs:                       2383.0
Min Obs:                       2383.0
Max Obs:                       2383.0

                              Parameter Estimates
===============================================================================
             Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
-------------------------------------------------------------------------------
temp_pop        0.0138     0.0043     3.1721     0.0015      0.0053      0.0223
precip_pop     -0.0033     0.0081    -0.4103     0.6816     -0.0191      0.0125
temp_income     0.0028     0.0018     1.5827     0.1135     -0.0007      0.0062
===============================================================================

F-test for Poolability: 101.78
P-value: 0.0000
Distribution: F(2585,483543)

Included effects: Entity, Time
```

In [33]:

```
#TABLE 4 COLUMN 6
table_4_gini_df = muni[muni['year'] <= 2006].dropna(subset = ['hom_rate', 'temp_pop
table_4_gini_df = table_4_gini_df.set_index(['muni_code', 'date'])
table_4_gini_df = table_4_gini_df[['state', 'hom_rate', 'popw', 'temp_pop', 'precip
table_4_gini_df
```

Out[33]:

muni_code	date	state	hom_rate	popw	temp_pop	precip_pop	temp_gin
1001	2001-04-01	Aguascalientes	0.301511	655030.9375	23.817400	0.113894	2.16750
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	0.000000	1.43586
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	0.038703	-1.23313
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	0.002461	-0.61166
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	1.110245	1.27620
...	...	...	...	...	...	...	.
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	1.465000	-0.38565
	1993-03-01	Zacatecas	2.608534	124968.4375	13.100000	0.016000	0.16496
	2003-07-01	Zacatecas	0.000000	124968.4375	16.400000	1.907000	-0.06216
	2005-12-01	Zacatecas	1.503913	124968.4375	11.900000	0.031000	0.24756
	2004-11-01	Zacatecas	1.517868	124968.4375	13.100000	0.029000	0.16496

482868 rows × 6 columns

In [34]:

```
mod_gini = PanelOLS.from_formula('hom_rate ~ temp_pop + precip_pop + temp_gini + Ti

result_gini = mod_gini.fit(cov_type='clustered', clusters = table_4_gini_df['state'

print(result_gini.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1638
No. Observations:              482868   R-squared (Within):              -0.0001
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0741
Time:                        15:13:20   Log-likelihood                -1.026e+06
Cov. Estimator:             Clustered
                                        F-statistic:                      32.459
Entities:                        2367   P-value                           0.0000
Avg Obs:                       204.00   Distribution:                F(3,480295)
Min Obs:                       204.00
Max Obs:                       204.00   F-statistic (robust):             6.7431
                                        P-value                           0.0002
Time periods:                     204   Distribution:                F(3,480295)
Avg Obs:                       2367.0
Min Obs:                       2367.0
Max Obs:                       2367.0

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0155     0.0037     4.1984     0.0000      0.0083      0.0228
precip_pop    -0.0036     0.0082    -0.4453     0.6561     -0.0196      0.0124
temp_gini      0.0008     0.0026     0.3126     0.7546     -0.0043      0.0059
==============================================================================

F-test for Poolability: 102.24
P-value: 0.0000
Distribution: F(2569,480295)

Included effects: Entity, Time
```

In [35]:
```python
#TABLE 4 COLUMN 7
table_4_ac_df = muni[(muni['year'] <= 2006) & (muni['ac_data'] == 1)].dropna(subset
table_4_ac_df = table_4_ac_df.set_index(['muni_code', 'date'])
table_4_ac_df = table_4_ac_df[['state', 'hom_rate', 'popw', 'temp_pop', 'precip_pop
table_4_ac_df
```

Out[35]:

muni_code	date	state	hom_rate	popw	temp_pop	precip_pop	temp_ac
1001	2001-04-01	Aguascalientes	0.301511	655030.9375	23.817400	0.113894	-1.352743
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	0.000000	-0.896125
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	0.038703	0.769604
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	0.002461	0.381742
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	1.110245	-0.796479
...	...	...	...	...	...	...	...
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	1.465000	-2.459257
	1993-03-01	Zacatecas	2.608534	124968.4375	13.100000	0.016000	1.051976
	2003-07-01	Zacatecas	0.000000	124968.4375	16.400000	1.907000	-0.396407
	2005-12-01	Zacatecas	1.503913	124968.4375	11.900000	0.031000	1.578662
	2004-11-01	Zacatecas	1.517868	124968.4375	13.100000	0.029000	1.051976

121056 rows × 6 columns

In [36]:

```
mod_ac = PanelOLS.from_formula('hom_rate ~ temp_pop + precip_pop + temp_ac + TimeEf
result_hom = mod_ac.fit(cov_type='clustered', clusters = table_4_ac_df['state'])
print(result_hom.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0005
Estimator:                   PanelOLS   R-squared (Between):              0.2073
No. Observations:              121056   R-squared (Within):              -0.0003
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.1263
Time:                        15:13:33   Log-likelihood                -2.193e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      20.628
Entities:                         599   P-value                           0.0000
Avg Obs:                       202.10   Distribution:                F(3,120251)
Min Obs:                       24.000
Max Obs:                       204.00   F-statistic (robust):             7.9397
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(3,120251)
Avg Obs:                       593.41
Min Obs:                       585.00
Max Obs:                       599.00

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0218     0.0053     4.1176     0.0000      0.0114      0.0321
precip_pop    -0.0049     0.0089    -0.5526     0.5805     -0.0223      0.0125
temp_ac       -0.0022     0.0012    -1.8150     0.0695     -0.0045      0.0002
==============================================================================

F-test for Poolability: 154.13
P-value: 0.0000
Distribution: F(801,120251)

Included effects: Entity, Time
```

In [37]:

```
#TABLE 4 COLUMN 8
table_4_at_df = muni[muni['year'] <= 2006].dropna(subset = ['hom_rate', 'temp_pop',
table_4_at_df = table_4_at_df.set_index(['muni_code', 'date'])
table_4_at_df = table_4_at_df[['state', 'hom_rate', 'popw', 'temp_pop', 'precip_pop
table_4_at_df
```

Out[37]:

muni_code	date	state	hom_rate	popw	temp_pop	precip_pop	temp_av
1001	2001-04-01	Aguascalientes	0.301511	655030.9375	23.817400	0.113894	-1.
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	0.000000	-0.
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	0.038703	0.
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	0.002461	0.
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	1.110245	-0.
...	...	...	...	...	...	...	
32058	2006-11-01	Zacatecas	0.000000	2711.0000	18.696400	0.192440	-1.
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000	-1.
	2005-03-01	Zacatecas	38.355366	2711.0000	19.163700	0.044328	-0.
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.002020	-0.
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.608296	1.

493908 rows × 6 columns

In [38]:

```
mod_at= PanelOLS.from_formula('hom_rate ~ temp_pop + precip_pop + temp_avg_temp + T

result_at = mod_at.fit(cov_type='clustered', clusters = table_4_at_df['state'])

print(result_at.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1660
No. Observations:              493908   R-squared (Within):              -0.0002
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0748
Time:                        15:15:03   Log-likelihood                -1.051e+06
Cov. Estimator:             Clustered
                                        F-statistic:                      32.758
Entities:                        2447   P-value                           0.0000
Avg Obs:                       201.84   Distribution:                F(3,491255)
Min Obs:                       24.000
Max Obs:                       204.00   F-statistic (robust):             6.1980
                                        P-value                           0.0003
Time periods:                     204   Distribution:                F(3,491255)
Avg Obs:                       2421.1
Min Obs:                       2396.0
Max Obs:                       2447.0

                               Parameter Estimates
=================================================================================
               Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
---------------------------------------------------------------------------------
temp_pop          0.0157     0.0038     4.0986     0.0000      0.0082      0.0232
precip_pop       -0.0031     0.0082    -0.3729     0.7093     -0.0192      0.0130
temp_avg_temp     0.0003     0.0018     0.1522     0.8790     -0.0032      0.0038
=================================================================================

F-test for Poolability: 100.44
P-value: 0.0000
Distribution: F(2649,491255)

Included effects: Entity, Time
```

Table 5: Progresa Transfers

In [39]:

```
#TABLE 5 COLUMN 4
table_5_temp_df = muni[(muni['year'] <= 2006) & (muni['year'] >= 1998)].dropna(subs
table_5_temp_df = table_5_temp_df.set_index(['muni_code', 'date'])
table_5_temp_df = table_5_temp_df[['state', 'hom_rate', 'popw', 'temp_pop', 'precip
table_5_temp_df
```

Out[39]:

muni_code	date	state	hom_rate	popw	temp_pop	precip_pop
1001	**2001-04-01**	Aguascalientes	0.301511	655030.9375	23.817400	0.113894
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	0.000000
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	0.038703
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	0.002461
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	1.110245
...	**...**	...	...	...	...	...
32058	**2006-11-01**	Zacatecas	0.000000	2711.0000	18.696400	0.192440
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000
	2005-03-01	Zacatecas	38.355366	2711.0000	19.163700	0.044328
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.002020
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.608296

262992 rows × 5 columns

In [40]:

```
mod_temp = PanelOLS.from_formula('hom_rate ~ temp_pop + precip_pop + TimeEffects +

result_temp = mod_temp.fit(cov_type='clustered', clusters = table_5_temp_df['state'

print(result_temp.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1777
No. Observations:              262992   R-squared (Within):              -0.0004
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0794
Time:                        15:15:27   Log-likelihood                -4.977e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      23.621
Entities:                        2447   P-value                           0.0000
Avg Obs:                       107.48   Distribution:                F(2,260436)
Min Obs:                       24.000
Max Obs:                       108.00   F-statistic (robust):             8.2345
                                        P-value                           0.0003
Time periods:                     108   Distribution:                F(2,260436)
Avg Obs:                       2435.1
Min Obs:                       2421.0
Max Obs:                       2447.0

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0127     0.0032     4.0075     0.0001      0.0065      0.0190
precip_pop     0.0071     0.0084     0.8434     0.3990     -0.0094      0.0235
==============================================================================

F-test for Poolability: 51.954
P-value: 0.0000
Distribution: F(2553,260436)

Included effects: Entity, Time
```

In [41]:

```
#TABLE 5 COLUMN 5
table_5_pro_df = muni[(muni['year'] <= 2006) & (muni['year'] >= 1998)].dropna(subse
table_5_pro_df = table_5_pro_df.set_index(['muni_code', 'date'])
table_5_pro_df = table_5_pro_df[['state', 'hom_rate', 'popw', 'temp_pop', 'progresa
table_5_pro_df
```

Out[41]:

muni_code	date	state	hom_rate	popw	temp_pop	progresa_per_capita
1001	2001-04-01	Aguascalientes	0.301511	655030.9375	23.817400	0.002714
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	0.000000
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	0.004456
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	0.002706
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	0.001304
...	...	...	...	...	...	...
32058	2006-11-01	Zacatecas	0.000000	2711.0000	18.696400	0.000000
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000
	2005-03-01	Zacatecas	38.355366	2711.0000	19.163700	0.000000
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.000000
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.000000

262992 rows × 6 columns

In [42]:

```
mod_pro = PanelOLS.from_formula('hom_rate ~ temp_pop + precip_pop + progresa_per_ca

result_pro = mod_pro.fit(cov_type='clustered', clusters = table_5_pro_df['state'])

print(result_pro.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1778
No. Observations:              262992   R-squared (Within):              -0.0004
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0795
Time:                        15:15:46   Log-likelihood                -4.977e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      15.805
Entities:                        2447   P-value                           0.0000
Avg Obs:                       107.48   Distribution:                F(3,260435)
Min Obs:                       24.000
Max Obs:                       108.00   F-statistic (robust):             5.5385
                                        P-value                           0.0008
Time periods:                     108   Distribution:                F(3,260435)
Avg Obs:                       2435.1
Min Obs:                       2421.0
Max Obs:                       2447.0

                                Parameter Estimates
====================================================================================
===
                      Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper
CI
------------------------------------------------------------------------------------
---
temp_pop                 0.0127     0.0032     4.0060     0.0001      0.0065       0.0
190
precip_pop               0.0071     0.0084     0.8449     0.3982     -0.0093       0.0
235
progresa_per_capita      0.0042     0.0114     0.3662     0.7142     -0.0182       0.0
266
====================================================================================
===

F-test for Poolability: 51.938
P-value: 0.0000
Distribution: F(2553,260435)

Included effects: Entity, Time
```

In [43]:

```python
#TABLE 5 COLUMN 6
table_5_pt_df = muni[(muni['year'] <= 2006) & (muni['year'] >= 1998)].dropna(subset
table_5_pt_df = table_5_pt_df.set_index(['muni_code', 'date'])
table_5_pt_df = table_5_pt_df[['state', 'hom_rate', 'popw', 'temp_pop', 'progresa_p
table_5_pt_df
```

Out[43]:

muni_code	date	state	hom_rate	popw	temp_pop	progresa_per_capita
1001	2001-04-01	Aguascalientes	0.301511	655030.9375	23.817400	0.002714
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	0.000000
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	0.004456
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	0.002706
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	0.001304
...	...	...	...	...	...	...
32058	2006-11-01	Zacatecas	0.000000	2711.0000	18.696400	0.000000
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000
	2005-03-01	Zacatecas	38.355366	2711.0000	19.163700	0.000000
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.000000
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.000000

262992 rows × 7 columns

In [44]:

```
mod_pt = PanelOLS.from_formula('hom_rate ~ temp_pop + precip_pop + progresa_per_cap

result_pt = mod_pt.fit(cov_type='clustered', clusters = table_5_pt_df['state'])

print(result_pt.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1780
No. Observations:              262992   R-squared (Within):              -0.0004
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0795
Time:                        15:16:03   Log-likelihood                -4.977e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      11.951
Entities:                        2447   P-value                           0.0000
Avg Obs:                       107.48   Distribution:                F(4,260434)
Min Obs:                       24.000
Max Obs:                       108.00   F-statistic (robust):             4.1933
                                        P-value                           0.0021
Time periods:                     108   Distribution:                F(4,260434)
Avg Obs:                       2435.1
Min Obs:                       2421.0
Max Obs:                       2447.0

                                Parameter Estimates
====================================================================================
===
                     Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper
CI
------------------------------------------------------------------------------------
---
temp_pop                0.0128     0.0032     4.0042     0.0001      0.0065      0.0
190
precip_pop              0.0071     0.0084     0.8483     0.3963     -0.0093      0.0
235
progresa_per_capita     0.0045     0.0116     0.3866     0.6991     -0.0183      0.0
272
temp_progresa          -0.0024     0.0024    -1.0216     0.3070     -0.0070      0.0
022
====================================================================================
===

F-test for Poolability: 51.938
P-value: 0.0000
Distribution: F(2553,260434)

Included effects: Entity, Time
```

Table 6: Interaction with Agricultural Variables

In [45]:

```
#Column 5
table_6_hom_df = muni[muni['year'] <= 2006].dropna(subset = ['hom_rate', 'temp_pop'
table_6_hom_df = table_6_hom_df.set_index(['muni_code', 'date'])
table_6_hom_df = table_6_hom_df[['state', 'hom_rate', 'popw', 'temp_pop', 'precip_p
table_6_hom_df
```

Out[45]:

		state	hom_rate	popw	temp_pop	precip_pop
muni_code	**date**					
1001	**2001-04-01**	Aguascalientes	0.301511	655030.9375	23.817400	0.113894
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	0.000000
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	0.038703
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	0.002461
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	1.110245
...	**...**	...	...	...	...	...
32058	**2006-11-01**	Zacatecas	0.000000	2711.0000	18.696400	0.192440
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000
	2005-03-01	Zacatecas	38.355366	2711.0000	19.163700	0.044328
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.002020
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.608296

493908 rows × 5 columns

In [46]:

```
mod_table_6 = PanelOLS.from_formula('hom_rate ~ temp_pop + precip_pop + TimeEffects

result_hom = mod_table_6.fit(cov_type='clustered', clusters = table_6_hom_df['state

print(result_hom.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1666
No. Observations:              493908   R-squared (Within):              -0.0002
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0750
Time:                        15:17:34   Log-likelihood                -1.051e+06
Cov. Estimator:             Clustered
                                        F-statistic:                      49.099
Entities:                        2447   P-value                           0.0000
Avg Obs:                       201.84   Distribution:                F(2,491256)
Min Obs:                       24.000
Max Obs:                       204.00   F-statistic (robust):             9.2778
                                        P-value                           0.0001
Time periods:                     204   Distribution:                F(2,491256)
Avg Obs:                       2421.1
Min Obs:                       2396.0
Max Obs:                       2447.0

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0158     0.0039     4.0921     0.0000      0.0082      0.0233
precip_pop    -0.0031     0.0082    -0.3758     0.7071     -0.0191      0.0130
==============================================================================

F-test for Poolability: 100.44
P-value: 0.0000
Distribution: F(2649,491256)

Included effects: Entity, Time
```

In [47]:

```
#Column 6
table_6_hom_df_growing = muni[muni['year'] <= 2006].dropna(subset = ['hom_rate', 't
table_6_hom_df_growing = table_6_hom_df_growing.set_index(['muni_code', 'date'])
table_6_hom_df_growing = table_6_hom_df_growing[['state', 'hom_rate', 'popw', 'temp
table_6_hom_df_growing
```

Out[47]:

		state	hom_rate	popw	temp_pop	precip_pop	temp_gr
muni_code	**date**						
1001	**2001-04-01**	Aguascalientes	0.301511	655030.9375	23.817400	0.113894	
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	0.000000	
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	0.038703	
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	0.002461	
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	1.110245	
...	**...**	...	...	...	...	...	
32058	**2006-11-01**	Zacatecas	0.000000	2711.0000	18.696400	0.192440	
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000	
	2005-03-01	Zacatecas	38.355366	2711.0000	19.163700	0.044328	
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.002020	
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.608296	

493908 rows × 7 columns

In [48]:

```
mod_table_6_growing = PanelOLS.from_formula('hom_rate ~ temp_pop + temp_growing_sea

result_hom_growing = mod_table_6_growing.fit(cov_type='clustered', clusters = table

print(result_hom_growing.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1449
No. Observations:              493908   R-squared (Within):              -0.0001
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0653
Time:                        15:19:06   Log-likelihood                -1.051e+06
Cov. Estimator:             Clustered
                                        F-statistic:                      33.119
Entities:                        2447   P-value                           0.0000
Avg Obs:                       201.84   Distribution:                F(3,491255)
Min Obs:                       24.000
Max Obs:                       204.00   F-statistic (robust):             7.1735
                                        P-value                           0.0001
Time periods:                     204   Distribution:                F(3,491255)
Avg Obs:                       2421.1
Min Obs:                       2396.0
Max Obs:                       2447.0

                                Parameter Estimates
===================================================================================
===
                     Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper
CI
-----------------------------------------------------------------------------------
---
temp_pop                0.0131     0.0051     2.5865     0.0097      0.0032      0.0
231
temp_growing_season     0.0050     0.0065     0.7715     0.4404     -0.0078      0.0
178
precip_pop             -0.0028     0.0084    -0.3276     0.7432     -0.0193      0.0
138
===================================================================================
===

F-test for Poolability: 100.40
P-value: 0.0000
Distribution: F(2649,491255)

Included effects: Entity, Time
```

In [49]:

```
#Column 7
table_6_hom_df_rural = muni[muni['year'] <= 2006].dropna(subset = ['hom_rate', 'tem
table_6_hom_df_rural = table_6_hom_df_rural.set_index(['muni_code', 'date'])
table_6_hom_df_rural = table_6_hom_df_rural[['state', 'hom_rate', 'popw', 'temp_pop
table_6_hom_df_rural
```

Out[49]:

muni_code	date	state	hom_rate	popw	temp_pop	temp_rural	precip_pc
1001	2001-04-01	Aguascalientes	0.301511	655030.9375	23.817400	-7.394476	0.11389
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	-4.898471	0.00000
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	4.206873	0.03870
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	2.086713	0.00246
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	-4.353780	1.11024
...	...	...	...	...	...	...	
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	-8.945066	1.46500
	1993-03-01	Zacatecas	2.608534	124968.4375	13.100000	3.826359	0.01600
	2003-07-01	Zacatecas	0.000000	124968.4375	16.400000	-1.441853	1.90700
	2005-12-01	Zacatecas	1.503913	124968.4375	11.900000	5.742074	0.03100
	2004-11-01	Zacatecas	1.517868	124968.4375	13.100000	3.826359	0.02900

488784 rows × 6 columns

In [50]:
```
mod_table_6_rural = PanelOLS.from_formula('hom_rate ~ temp_pop + temp_rural + preci

result_hom_rural = mod_table_6_rural.fit(cov_type='clustered', clusters = table_6_h

print(result_hom_rural.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1606
No. Observations:              488784   R-squared (Within):              -0.0001
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0724
Time:                        15:20:37   Log-likelihood                  -1.04e+06
Cov. Estimator:             Clustered
                                        F-statistic:                      32.617
Entities:                        2396   P-value                           0.0000
Avg Obs:                       204.00   Distribution:                F(3,486182)
Min Obs:                       204.00
Max Obs:                       204.00   F-statistic (robust):             7.2288
                                        P-value                           0.0001
Time periods:                     204   Distribution:                F(3,486182)
Avg Obs:                       2396.0
Min Obs:                       2396.0
Max Obs:                       2396.0

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0152     0.0055     2.7756     0.0055      0.0045      0.0259
temp_rural    -0.0004     0.0018    -0.2185     0.8270     -0.0040      0.0032
precip_pop    -0.0036     0.0082    -0.4471     0.6548     -0.0196      0.0123
==============================================================================

F-test for Poolability: 101.68
P-value: 0.0000
Distribution: F(2598,486182)

Included effects: Entity, Time
```

In [51]:

```
#Column 8
table_6_hom_df_agric = muni[muni['year'] <= 2006].dropna(subset = ['hom_rate', 'tem
table_6_hom_df_agric = table_6_hom_df_agric.set_index(['muni_code', 'date'])
table_6_hom_df_agric = table_6_hom_df_agric[['state', 'hom_rate', 'popw', 'temp_pop
table_6_hom_df_agric
```

Out[51]:

muni_code	date	state	hom_rate	popw	temp_pop	temp_agric	precip_p
1001	2001-04-01	Aguascalientes	0.301511	655030.9375	23.817400	-9.399092	0.1138
	1998-04-01	Aguascalientes	0.642868	655030.9375	22.207600	-6.226429	0.0000
	2001-11-01	Aguascalientes	0.297347	655030.9375	16.335100	5.347341	0.0387
	2000-02-01	Aguascalientes	0.000000	655030.9375	17.702499	2.652413	0.0024
	2001-07-01	Aguascalientes	0.149856	655030.9375	21.856300	-5.534073	1.1102
...	...	...	...	...	...	...	
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	-11.386839	1.4650
	1993-03-01	Zacatecas	2.608534	124968.4375	13.100000	4.870856	0.0160
	2003-07-01	Zacatecas	0.000000	124968.4375	16.400000	-1.835442	1.9070
	2005-12-01	Zacatecas	1.503913	124968.4375	11.900000	7.309511	0.0310
	2004-11-01	Zacatecas	1.517868	124968.4375	13.100000	4.870856	0.0290

488784 rows × 6 columns

In [52]:

```
mod_table_6_agric = PanelOLS.from_formula('hom_rate ~ temp_pop + temp_agric + preci
result_hom_agric = mod_table_6_agric.fit(cov_type='clustered', clusters = table_6_h
print(result_hom_agric.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               hom_rate   R-squared:                        0.0002
Estimator:                   PanelOLS   R-squared (Between):              0.1597
No. Observations:              488784   R-squared (Within):              -0.0001
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0721
Time:                        15:22:06   Log-likelihood                 -1.04e+06
Cov. Estimator:             Clustered
                                        F-statistic:                      32.632
Entities:                        2396   P-value                           0.0000
Avg Obs:                       204.00   Distribution:                F(3,486182)
Min Obs:                       204.00
Max Obs:                       204.00   F-statistic (robust):             7.2355
                                        P-value                           0.0001
Time periods:                     204   Distribution:                F(3,486182)
Avg Obs:                       2396.0
Min Obs:                       2396.0
Max Obs:                       2396.0

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0151     0.0055     2.7319     0.0063      0.0043      0.0259
temp_agric    -0.0004     0.0017    -0.2486     0.8037     -0.0036      0.0028
precip_pop    -0.0036     0.0082    -0.4446     0.6566     -0.0196      0.0124
==============================================================================

F-test for Poolability: 101.68
P-value: 0.0000
Distribution: F(2598,486182)

Included effects: Entity, Time
```

Table 8: Temperature and Suicides in Mexico

In [53]:
```
#Column 1
table_8_rural_df = muni[muni['year'] <= 2006].dropna(subset = ['sui_rate', 'temp_po
table_8_rural_df = table_8_rural_df.set_index(['muni_code', 'date'])
table_8_rural_df = table_8_rural_df[['state', 'sui_rate', 'popw', 'temp_pop', 'prec
table_8_rural_df
```

Out[53]:

muni_code	date	state	sui_rate	popw	temp_pop	precip_pop
1001	2001-04-01	Aguascalientes	0.150756	655030.9375	23.817400	0.113894
	1998-04-01	Aguascalientes	0.160717	655030.9375	22.207600	0.000000
	2001-11-01	Aguascalientes	0.446021	655030.9375	16.335100	0.038703
	2000-02-01	Aguascalientes	0.620399	655030.9375	17.702499	0.002461
	2001-07-01	Aguascalientes	0.449569	655030.9375	21.856300	1.110245
...	...	...	...	...	...	...
32058	2006-11-01	Zacatecas	0.000000	2711.0000	18.696400	0.192440
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000
	2005-03-01	Zacatecas	0.000000	2711.0000	19.163700	0.044328
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.002020
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.608296

493908 rows × 5 columns

In [54]:

```
mod_sui8 = PanelOLS.from_formula('sui_rate ~ temp_pop + precip_pop + TimeEffects +
result_sui = mod_sui8.fit(cov_type='clustered', clusters = table_8_rural_df['state'
print(result_sui.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               sui_rate   R-squared:                        0.0003
Estimator:                   PanelOLS   R-squared (Between):              0.2921
No. Observations:              493908   R-squared (Within):               0.0006
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0765
Time:                        15:23:39   Log-likelihood                -5.606e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      85.888
Entities:                        2447   P-value                           0.0000
Avg Obs:                       201.84   Distribution:                F(2,491256)
Min Obs:                       24.000
Max Obs:                       204.00   F-statistic (robust):             33.552
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(2,491256)
Avg Obs:                       2421.1
Min Obs:                       2396.0
Max Obs:                       2447.0

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0072     0.0010     7.2590     0.0000      0.0052      0.0091
precip_pop    -0.0048     0.0024    -2.0513     0.0402     -0.0094     -0.0002
==============================================================================

F-test for Poolability: 40.201
P-value: 0.0000
Distribution: F(2649,491256)

Included effects: Entity, Time
```

In [55]:

```
#Column 2
table_8_rural_df = muni[muni['year'] <= 2006].dropna(subset = ['sui_rate', 'temp_po
table_8_rural_df = table_8_rural_df.set_index(['muni_code', 'date'])
table_8_rural_df = table_8_rural_df[['state', 'sui_rate', 'popw', 'temp_pop', 'prec
table_8_rural_df
```

Out[55]:

muni_code	date	state	sui_rate	popw	temp_pop	precip_pop	temp_incc
1001	2001-04-01	Aguascalientes	0.150756	655030.9375	23.817400	0.113894	2.531
	1998-04-01	Aguascalientes	0.160717	655030.9375	22.207600	0.000000	1.676
	2001-11-01	Aguascalientes	0.446021	655030.9375	16.335100	0.038703	-1.440
	2000-02-01	Aguascalientes	0.620399	655030.9375	17.702499	0.002461	-0.714
	2001-07-01	Aguascalientes	0.449569	655030.9375	21.856300	1.110245	1.490
...	...	...	...	...	...	...	
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	1.465000	1.397
	1993-03-01	Zacatecas	0.869511	124968.4375	13.100000	0.016000	-0.597
	2003-07-01	Zacatecas	0.771639	124968.4375	16.400000	1.907000	0.225
	2005-12-01	Zacatecas	0.751956	124968.4375	11.900000	0.031000	-0.896
	2004-11-01	Zacatecas	0.758934	124968.4375	13.100000	0.029000	-0.597

486132 rows × 6 columns

In [56]:

```
mod_inc = PanelOLS.from_formula('sui_rate ~ temp_pop + precip_pop + temp_income + T

result_inc8 = mod_inc.fit(cov_type='clustered', clusters = table_8_rural_df['state'

print(result_inc8.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               sui_rate   R-squared:                        0.0004
Estimator:                   PanelOLS   R-squared (Between):              0.2708
No. Observations:              486132   R-squared (Within):               0.0006
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0716
Time:                        15:25:16   Log-likelihood                -5.497e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      58.285
Entities:                        2383   P-value                           0.0000
Avg Obs:                       204.00   Distribution:                F(3,483543)
Min Obs:                       204.00
Max Obs:                       204.00   F-statistic (robust):             22.350
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(3,483543)
Avg Obs:                       2383.0
Min Obs:                       2383.0
Max Obs:                       2383.0

                              Parameter Estimates
===============================================================================
             Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
-------------------------------------------------------------------------------
temp_pop        0.0065     0.0010     6.3719     0.0000      0.0045      0.0085
precip_pop     -0.0048     0.0024    -1.9550     0.0506     -0.0095   1.217e-05
temp_income     0.0010     0.0009     1.1492     0.2505     -0.0007      0.0028
===============================================================================

F-test for Poolability: 41.090
P-value: 0.0000
Distribution: F(2585,483543)

Included effects: Entity, Time
```

In [57]:

```
#Column 3
table_8_gini_df = muni[muni['year'] <= 2006].dropna(subset = ['sui_rate', 'temp_pop
table_8_gini_df = table_8_gini_df.set_index(['muni_code', 'date'])
table_8_gini_df = table_8_gini_df[['state', 'sui_rate', 'popw', 'temp_pop', 'precip
table_8_gini_df
```

Out[57]:

muni_code	date	state	sui_rate	popw	temp_pop	precip_pop	temp_gini
1001	2001-04-01	Aguascalientes	0.150756	655030.9375	23.817400	0.113894	2.167502
	1998-04-01	Aguascalientes	0.160717	655030.9375	22.207600	0.000000	1.435862
	2001-11-01	Aguascalientes	0.446021	655030.9375	16.335100	0.038703	-1.233138
	2000-02-01	Aguascalientes	0.620399	655030.9375	17.702499	0.002461	-0.611667
	2001-07-01	Aguascalientes	0.449569	655030.9375	21.856300	1.110245	1.276200
...	...	...	...	...	...	...	...
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	1.465000	-0.385651
	1993-03-01	Zacatecas	0.869511	124968.4375	13.100000	0.016000	0.164967
	2003-07-01	Zacatecas	0.771639	124968.4375	16.400000	1.907000	-0.062163
	2005-12-01	Zacatecas	0.751956	124968.4375	11.900000	0.031000	0.247560
	2004-11-01	Zacatecas	0.758934	124968.4375	13.100000	0.029000	0.164967

482868 rows × 6 columns

In [58]:

```
mod_gini8 = PanelOLS.from_formula('sui_rate ~ temp_pop + precip_pop + temp_gini + T

result_gini8 = mod_gini8.fit(cov_type='clustered', clusters = table_8_gini_df['stat

print(result_gini8.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               sui_rate   R-squared:                        0.0004
Estimator:                   PanelOLS   R-squared (Between):              0.2886
No. Observations:              482868   R-squared (Within):               0.0006
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0764
Time:                        15:26:46   Log-likelihood                -5.452e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      57.417
Entities:                        2367   P-value                           0.0000
Avg Obs:                       204.00   Distribution:                F(3,480295)
Min Obs:                       204.00
Max Obs:                       204.00   F-statistic (robust):             23.819
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(3,480295)
Avg Obs:                       2367.0
Min Obs:                       2367.0
Max Obs:                       2367.0

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0071     0.0010     7.3739     0.0000      0.0052      0.0089
precip_pop    -0.0048     0.0024    -2.0124     0.0442     -0.0096     -0.0001
temp_gini      0.0005     0.0014     0.3476     0.7281     -0.0023      0.0033
==============================================================================

F-test for Poolability: 41.198
P-value: 0.0000
Distribution: F(2569,480295)

Included effects: Entity, Time
```

In [59]:

```
#Column 4
table_8_ac_df = muni[(muni['year'] <= 2006) & (muni['ac_data'] == 1)].dropna(subset
table_8_ac_df = table_8_ac_df.set_index(['muni_code', 'date'])
table_8_ac_df = table_8_ac_df[['state', 'sui_rate', 'popw', 'temp_pop', 'precip_pop
table_8_ac_df
```

Out[59]:

muni_code	date	state	sui_rate	popw	temp_pop	precip_pop	temp_ac
1001	2001-04-01	Aguascalientes	0.150756	655030.9375	23.817400	0.113894	-1.352743
	1998-04-01	Aguascalientes	0.160717	655030.9375	22.207600	0.000000	-0.896125
	2001-11-01	Aguascalientes	0.446021	655030.9375	16.335100	0.038703	0.769604
	2000-02-01	Aguascalientes	0.620399	655030.9375	17.702499	0.002461	0.381742
	2001-07-01	Aguascalientes	0.449569	655030.9375	21.856300	1.110245	-0.796479
...	...	...	...	...	...	...	...
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	1.465000	-2.459257
	1993-03-01	Zacatecas	0.869511	124968.4375	13.100000	0.016000	1.051976
	2003-07-01	Zacatecas	0.771639	124968.4375	16.400000	1.907000	-0.396407
	2005-12-01	Zacatecas	0.751956	124968.4375	11.900000	0.031000	1.578662
	2004-11-01	Zacatecas	0.758934	124968.4375	13.100000	0.029000	1.051976

121056 rows × 6 columns

In [60]:

```
mod_ac8 = PanelOLS.from_formula('sui_rate ~ temp_pop + precip_pop + temp_ac + TimeE
result_ac8 = mod_ac8.fit(cov_type='clustered', clusters = table_8_ac_df['state'])
print(result_ac8.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               sui_rate   R-squared:                        0.0009
Estimator:                   PanelOLS   R-squared (Between):              0.1955
No. Observations:              121056   R-squared (Within):               0.0015
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0922
Time:                        15:27:02   Log-likelihood                -9.227e+04
Cov. Estimator:             Clustered
                                        F-statistic:                      36.915
Entities:                         599   P-value                           0.0000
Avg Obs:                       202.10   Distribution:                F(3,120251)
Min Obs:                       24.000
Max Obs:                       204.00   F-statistic (robust):             37.087
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(3,120251)
Avg Obs:                       593.41
Min Obs:                       585.00
Max Obs:                       599.00

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0049     0.0019     2.6588     0.0078      0.0013      0.0086
precip_pop    -0.0059     0.0025    -2.3521     0.0187     -0.0108     -0.0010
temp_ac        0.0009     0.0009     0.9506     0.3418     -0.0010      0.0027
==============================================================================

F-test for Poolability: 83.472
P-value: 0.0000
Distribution: F(801,120251)

Included effects: Entity, Time
```

In [61]:

```
# Column 5
table_8_at_df = muni[muni['year'] <= 2006].dropna(subset = ['sui_rate', 'temp_pop',
table_8_at_df = table_8_at_df.set_index(['muni_code', 'date'])
table_8_at_df = table_8_at_df[['state', 'sui_rate', 'popw', 'temp_pop', 'precip_pop
table_8_at_df
```

Out[61]:

muni_code	date	state	sui_rate	popw	temp_pop	precip_pop	temp_avg
1001	2001-04-01	Aguascalientes	0.150756	655030.9375	23.817400	0.113894	-1.2
	1998-04-01	Aguascalientes	0.160717	655030.9375	22.207600	0.000000	-0.7
	2001-11-01	Aguascalientes	0.446021	655030.9375	16.335100	0.038703	0.6
	2000-02-01	Aguascalientes	0.620399	655030.9375	17.702499	0.002461	0.3
	2001-07-01	Aguascalientes	0.449569	655030.9375	21.856300	1.110245	-0.7
...	...	...	...	...	...	...	
32058	2006-11-01	Zacatecas	0.000000	2711.0000	18.696400	0.192440	-1.1
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000	-1.6
	2005-03-01	Zacatecas	0.000000	2711.0000	19.163700	0.044328	-0.9
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.002020	-0.2
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.608296	1.9

493908 rows × 6 columns

In [62]:

```
mod_at8 = PanelOLS.from_formula('sui_rate ~ temp_pop + precip_pop + temp_avg_temp +

result_at8 = mod_at8.fit(cov_type='clustered', clusters = table_8_at_df['state'])

print(result_at8.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               sui_rate   R-squared:                        0.0004
Estimator:                   PanelOLS   R-squared (Between):              0.2840
No. Observations:              493908   R-squared (Within):               0.0006
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0744
Time:                        15:28:53   Log-likelihood                -5.606e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      60.651
Entities:                        2447   P-value                           0.0000
Avg Obs:                       201.84   Distribution:                F(3,491255)
Min Obs:                       24.000
Max Obs:                       204.00   F-statistic (robust):             22.507
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(3,491255)
Avg Obs:                       2421.1
Min Obs:                       2396.0
Max Obs:                       2447.0

                              Parameter Estimates
=================================================================================
               Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
---------------------------------------------------------------------------------
temp_pop          0.0069     0.0011     6.4216     0.0000      0.0048      0.0090
precip_pop       -0.0048     0.0023    -2.0692     0.0385     -0.0093     -0.0003
temp_avg_temp     0.0012     0.0007     1.7988     0.0721     -0.0001      0.0025
=================================================================================

F-test for Poolability: 40.200
P-value: 0.0000
Distribution: F(2649,491255)

Included effects: Entity, Time
```

In [63]:

```python
# Column 6
table_8_pt_df = muni[(muni['year'] <= 2006) & (muni['year'] >= 1998)].dropna(subset
table_8_pt_df = table_8_pt_df.set_index(['muni_code', 'date'])
table_8_pt_df = table_8_pt_df[['state', 'sui_rate', 'popw', 'temp_pop', 'precip_pop
table_8_pt_df
```

Out[63]:

muni_code	date	state	sui_rate	popw	temp_pop	precip_pop	progresa_
1001	2001-04-01	Aguascalientes	0.150756	655030.9375	23.817400	0.113894	
	1998-04-01	Aguascalientes	0.160717	655030.9375	22.207600	0.000000	
	2001-11-01	Aguascalientes	0.446021	655030.9375	16.335100	0.038703	
	2000-02-01	Aguascalientes	0.620399	655030.9375	17.702499	0.002461	
	2001-07-01	Aguascalientes	0.449569	655030.9375	21.856300	1.110245	
...	...	...	...	...	...	...	
32058	2006-11-01	Zacatecas	0.000000	2711.0000	18.696400	0.192440	
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000	
	2005-03-01	Zacatecas	0.000000	2711.0000	19.163700	0.044328	
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.002020	
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.608296	

262992 rows × 7 columns

In [64]:

```
mod_pt8 = PanelOLS.from_formula('sui_rate ~ temp_pop + precip_pop + progresa_per_ca

result_pt8 = mod_pt8.fit(cov_type='clustered', clusters = table_8_pt_df['state'])

print(result_pt8.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               sui_rate   R-squared:                        0.0005
Estimator:                   PanelOLS   R-squared (Between):              0.3705
No. Observations:              262992   R-squared (Within):               0.0007
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0987
Time:                        15:29:15   Log-likelihood                -3.103e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      33.831
Entities:                        2447   P-value                           0.0000
Avg Obs:                       107.48   Distribution:                F(4,260434)
Min Obs:                       24.000
Max Obs:                       108.00   F-statistic (robust):             13.204
                                        P-value                           0.0000
Time periods:                     108   Distribution:                F(4,260434)
Avg Obs:                       2435.1
Min Obs:                       2421.0
Max Obs:                       2447.0

                                Parameter Estimates
====================================================================================
===
                     Parameter  Std. Err.     T-stat    P-value    Lower CI     Upper
CI
------------------------------------------------------------------------------------
---
temp_pop                0.0093     0.0014     6.7722     0.0000      0.0066       0.0
120
precip_pop             -0.0044     0.0028    -1.6086     0.1077     -0.0099       0.0
010
progresa_per_capita    -0.0091     0.0050    -1.8277     0.0676     -0.0189       0.0
007
temp_progresa          -0.0020     0.0021    -0.9256     0.3546     -0.0061       0.0
022
====================================================================================
===

F-test for Poolability: 19.723
P-value: 0.0000
Distribution: F(2553,260434)

Included effects: Entity, Time
```

In [65]:

```
# Column 7
table_8_tg_df = muni[muni['year'] <= 2006].dropna(subset = ['sui_rate', 'temp_pop',
table_8_tg_df = table_8_tg_df.set_index(['muni_code', 'date'])
table_8_tg_df = table_8_tg_df[['state', 'sui_rate', 'popw', 'temp_pop', 'precip_pop
table_8_tg_df
```

Out[65]:

muni_code	date	state	sui_rate	popw	temp_pop	precip_pop	temp_gro
1001	2001-04-01	Aguascalientes	0.150756	655030.9375	23.817400	0.113894	
	1998-04-01	Aguascalientes	0.160717	655030.9375	22.207600	0.000000	
	2001-11-01	Aguascalientes	0.446021	655030.9375	16.335100	0.038703	
	2000-02-01	Aguascalientes	0.620399	655030.9375	17.702499	0.002461	
	2001-07-01	Aguascalientes	0.449569	655030.9375	21.856300	1.110245	
...	...	...	...	...	...	...	
32058	2006-11-01	Zacatecas	0.000000	2711.0000	18.696400	0.192440	
	2005-12-01	Zacatecas	0.000000	2711.0000	17.435801	0.000000	
	2005-03-01	Zacatecas	0.000000	2711.0000	19.163700	0.044328	
	2006-03-01	Zacatecas	0.000000	2711.0000	21.108900	0.002020	
	2005-06-01	Zacatecas	0.000000	2711.0000	26.785601	0.608296	

493908 rows × 7 columns

In [66]:

```
mod_tg = PanelOLS.from_formula('sui_rate ~ temp_pop + precip_pop + temp_growing_sea

result_tg = mod_tg.fit(cov_type='clustered', clusters = table_8_tg_df['state'])

print(result_tg.summary)
```

```
C:\Users\blahb\AppData\Local\Temp\ipykernel_16636\1795527819.py:3: AbsorbingEffectWa
rning:
Variables have been fully absorbed and have removed from the regression:

growing_season

  result_tg = mod_tg.fit(cov_type='clustered', clusters = table_8_tg_df['state'])
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               sui_rate   R-squared:                        0.0004
Estimator:                   PanelOLS   R-squared (Between):              0.3380
No. Observations:              493908   R-squared (Within):               0.0005
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0884
Time:                        15:31:09   Log-likelihood                -5.606e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      59.064
Entities:                        2447   P-value                           0.0000
Avg Obs:                       201.84   Distribution:                F(3,491255)
Min Obs:                       24.000
Max Obs:                       204.00   F-statistic (robust):             24.749
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(3,491255)
Avg Obs:                       2421.1
Min Obs:                       2396.0
Max Obs:                       2447.0

                              Parameter Estimates
====================================================================================
===
                    Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper 
CI
------------------------------------------------------------------------------------
---
temp_pop               0.0093     0.0017     5.4755     0.0000      0.0060      0.0
126
precip_pop            -0.0051     0.0023    -2.1731     0.0298     -0.0097     -0.0
005
temp_growing_season   -0.0040     0.0019    -2.1662     0.0303     -0.0077     -0.0
004
====================================================================================
===

F-test for Poolability: 40.147
P-value: 0.0000
Distribution: F(2649,491255)

Included effects: Entity, Time
```

In [67]:

```python
# Column 8
table_8_rural_df = muni[muni['year'] <= 2006].dropna(subset = ['sui_rate', 'temp_po
table_8_rural_df = table_8_rural_df.set_index(['muni_code', 'date'])
table_8_rural_df = table_8_rural_df[['state', 'sui_rate', 'popw', 'temp_pop', 'prec
table_8_rural_df
```

Out[67]:

muni_code	date	state	sui_rate	popw	temp_pop	precip_pop	temp_rura
1001	2001-04-01	Aguascalientes	0.150756	655030.9375	23.817400	0.113894	-7.39447
	1998-04-01	Aguascalientes	0.160717	655030.9375	22.207600	0.000000	-4.89847
	2001-11-01	Aguascalientes	0.446021	655030.9375	16.335100	0.038703	4.20687
	2000-02-01	Aguascalientes	0.620399	655030.9375	17.702499	0.002461	2.08671
	2001-07-01	Aguascalientes	0.449569	655030.9375	21.856300	1.110245	-4.35378
...	...	...	...	...	...	...	.
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	1.465000	-8.94506
	1993-03-01	Zacatecas	0.869511	124968.4375	13.100000	0.016000	3.82635
	2003-07-01	Zacatecas	0.771639	124968.4375	16.400000	1.907000	-1.44185
	2005-12-01	Zacatecas	0.751956	124968.4375	11.900000	0.031000	5.74207
	2004-11-01	Zacatecas	0.758934	124968.4375	13.100000	0.029000	3.82635

488784 rows × 6 columns

In [68]:

```
mod_rural = PanelOLS.from_formula('sui_rate ~ temp_pop + precip_pop + temp_rural +
result_rural = mod_rural.fit(cov_type='clustered', clusters = table_8_rural_df['sta
print(result_rural.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               sui_rate   R-squared:                        0.0004
Estimator:                   PanelOLS   R-squared (Between):              0.2801
No. Observations:              488784   R-squared (Within):               0.0006
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0739
Time:                        15:33:09   Log-likelihood                -5.531e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      57.850
Entities:                        2396   P-value                           0.0000
Avg Obs:                       204.00   Distribution:                F(3,486182)
Min Obs:                       204.00
Max Obs:                       204.00   F-statistic (robust):             23.182
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(3,486182)
Avg Obs:                       2396.0
Min Obs:                       2396.0
Max Obs:                       2396.0

                              Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0068     0.0015     4.6382     0.0000      0.0039      0.0096
precip_pop    -0.0049     0.0024    -2.0192     0.0435     -0.0096     -0.0001
temp_rural    -0.0003     0.0007    -0.4070     0.6840     -0.0017      0.0011
==============================================================================

F-test for Poolability: 41.028
P-value: 0.0000
Distribution: F(2598,486182)

Included effects: Entity, Time
```

In [69]:

```
#Column 9
table_8_agri_df = muni[muni['year'] <= 2006].dropna(subset = ['sui_rate', 'temp_pop
table_8_agri_df = table_8_agri_df.set_index(['muni_code', 'date'])
table_8_agri_df = table_8_agri_df[['state', 'sui_rate', 'popw', 'temp_pop', 'precip
table_8_agri_df
```

Out[69]:

muni_code	date	state	sui_rate	popw	temp_pop	precip_pop	temp_agri
1001	2001-04-01	Aguascalientes	0.150756	655030.9375	23.817400	0.113894	-9.39909
	1998-04-01	Aguascalientes	0.160717	655030.9375	22.207600	0.000000	-6.22642
	2001-11-01	Aguascalientes	0.446021	655030.9375	16.335100	0.038703	5.34734
	2000-02-01	Aguascalientes	0.620399	655030.9375	17.702499	0.002461	2.65241
	2001-07-01	Aguascalientes	0.449569	655030.9375	21.856300	1.110245	-5.53407
...	...	...	...	...	...	...	
32056	2003-05-01	Zacatecas	0.773258	124968.4375	21.100000	1.465000	-11.38683
	1993-03-01	Zacatecas	0.869511	124968.4375	13.100000	0.016000	4.87085
	2003-07-01	Zacatecas	0.771639	124968.4375	16.400000	1.907000	-1.83544
	2005-12-01	Zacatecas	0.751956	124968.4375	11.900000	0.031000	7.30951
	2004-11-01	Zacatecas	0.758934	124968.4375	13.100000	0.029000	4.87085

488784 rows × 6 columns

In [70]:

```
mod_agric = PanelOLS.from_formula('sui_rate ~ temp_pop + precip_pop + temp_agric + 
result_agric = mod_agric.fit(cov_type='clustered', clusters = table_8_agri_df['stat
print(result_agric.summary)
```

```
                          PanelOLS Estimation Summary
================================================================================
Dep. Variable:               sui_rate   R-squared:                        0.0004
Estimator:                   PanelOLS   R-squared (Between):              0.2695
No. Observations:              488784   R-squared (Within):               0.0006
Date:                Fri, Apr 26 2024   R-squared (Overall):              0.0711
Time:                        15:35:01   Log-likelihood                -5.531e+05
Cov. Estimator:             Clustered
                                        F-statistic:                      58.214
Entities:                        2396   P-value                           0.0000
Avg Obs:                       204.00   Distribution:                F(3,486182)
Min Obs:                       204.00
Max Obs:                       204.00   F-statistic (robust):             23.755
                                        P-value                           0.0000
Time periods:                     204   Distribution:                F(3,486182)
Avg Obs:                       2396.0
Min Obs:                       2396.0
Max Obs:                       2396.0

                             Parameter Estimates
==============================================================================
            Parameter  Std. Err.     T-stat    P-value    Lower CI    Upper CI
------------------------------------------------------------------------------
temp_pop       0.0064     0.0014     4.7461     0.0000      0.0038      0.0091
precip_pop    -0.0048     0.0024    -2.0001     0.0455     -0.0096  -9.685e-05
temp_agric    -0.0005     0.0006    -0.8419     0.3998     -0.0016      0.0006
==============================================================================

F-test for Poolability: 41.029
P-value: 0.0000
Distribution: F(2598,486182)

Included effects: Entity, Time
```