

Метод “Разность разностей”

Введение

В данной работе мы будем использовать результаты естественного эксперимента, в ходе которого рассматривалось влияние природного бедствия на уровень поддержки партии. В работе Бехтеля и Хайнмюллера (откуда и взяты данные) указывается, что наводнение на Эльбе произошло всего за месяц до федеральных выборов 2002 года и стало самым разрушительным наводнением в Германии за более чем тысячу лет. Это дало канцлеру Шредеру и его действующей Социал-демократической партии (СДПГ) ключевую возможность привлечь избирателей за счет массированных политических мер, чтобы предотвратить широко ожидаемое поражение на предстоящих выборах. В ответ на наводнение правительство быстро запустило крупнейшую программу помощи, которая и могла способствовать приросту доли голосов партии на предстоящих выборах.

Однако, как и авторы статьи мы хотим оценить не только краткосрочный эффект (который оказался сильным на выборах 2002 года), но рассмотреть долгосрочный эффект, например, на выборах 2005 года. Иными словами, мы хотим оценить, способствует ли помощь при природном бедствии поддержке партии не только на ближайших выборах, но и на последующих. И если да, то в какой мере сохраняется эффект – остается прежним, увеличивается или снижается? Для того, чтобы оценить это влияние, нам необходимо использовать метод “разность разностей”.

В ходе данной небольшой работы мы сначала обозначим данные, с которыми работаем, определим воздействующую переменную и зависимую, а также посмотрим на распределение экспериментальной и контрольной групп. Затем проверим допущение о параллельности трендов и оценим основной эффект на выборах 2005 года.

Данные

Единицами наблюдения являются избирательные округа Германии (всего их в работе 299 на 2005 год). Воздействие от наводнения в работе представлено бинарной переменной *Flooded* (1 - пострадал от наводнения, 0 – не пострадал). Округ считался пострадавшим при следующих ситуациях: стабилизация или прорыв дамб, предупреждение о наводнении, выход за пределы дамбы, затопление, предупреждение об эвакуации или эвакуация. Таким образом, в экспериментальной группе у нас 29 округов, а в контрольной – 270. В качестве опорных временных точек мы будем использовать t_1 – выборы 1998 года (до наводнения), t_2 – выборы 2005 года (спустя 3 года после наводнения).

Поддержка партии СДПГ операционализируется через долю голосов, отданных партии в конкретный год на выборах. Помимо основного эффекта, мы будем использовать дополнительно ряд ковариат: плотность населения (*popdens*), является ли СДПГ правящей партией в ландтаге соответствующей земли? (*SPD_inc*), уровень безработицы (*unemp_rate*), количество мигрантов на 1000 человек населения (*net.migration*).

Проверка допущений: параллельность трендов

Для проверки допущения о параллельных трендах, используется в качестве зависимой переменной значения до воздействия, т.е. до наводнения. В нашем случае это результаты выборов в 1994 году и в 1998 году. Мы проверяем, что при отсутствии

воздействия сохраняется параллельная тенденция - и для того, чтобы это проверить, мы используем обычную регрессию (функция *lm*) с использованием дамми-переменных.

На Рисунке 1 мы видим, что в среднем в экспериментальной группе (*Flooded*=1), по сравнению с контрольной группой (*Flooded*=0), поддержка СДПГ оставалась низкой в течении двух избирательных циклов. Исходя из этого наблюдения мы можем предположить, что допущение соблюдается, однако необходимо провести пару формальных тестов.

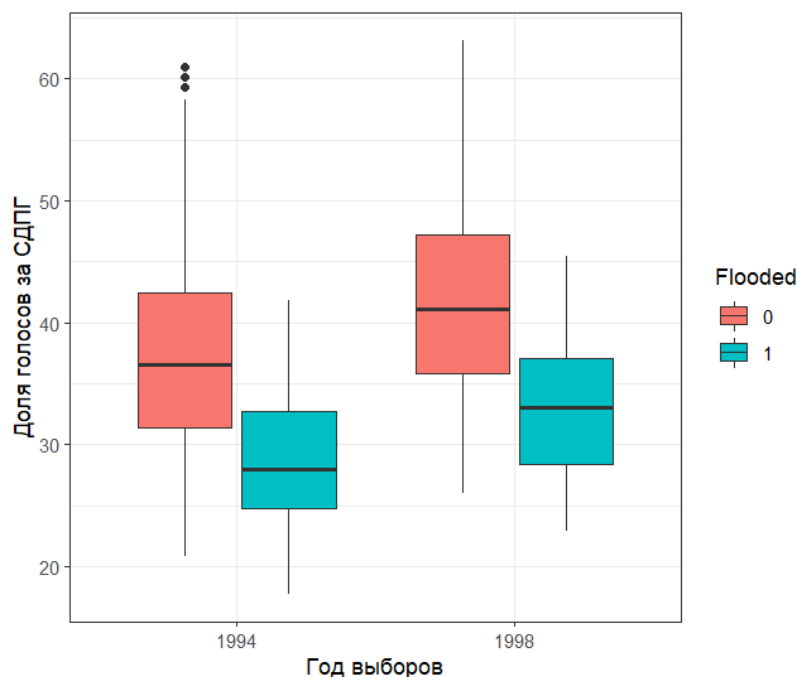


Рисунок 1. Диаграмма размаха доли голосов за СДПГ на выборах в 1994 году и 1998 году по двум группам округов: пострадавших во время наводнения 2002 года (*Flooded* = 1) и не пострадавших (*Flooded* = 0).

Проведя несколько тестов с различными спецификациями (см. Приложение 2-3), было обнаружено, что не просматривается эффекта воздействующей переменной *Flooded* ($p\text{-value}=0.999$). Следовательно, представляется вполне правдоподобным предположение, что в отсутствие наводнения экспериментальная группа и контрольная группа продолжала бы иметь относительно параллельные тренды в период после наводнения, так как доля голосов СДПГ в пострадавших и незатронутых округах имеет практически одинаковую тенденцию до наводнения на Эльбе.

Оценка эффекта

Для оценки эффекта использовалась обычная линейная регрессия (функция *lm*) с фиктивными переменными для электоральных округов. Согласно полученным результатам мы наблюдаем эффект от наводнения даже в 2005 году ($p\text{-value} = 0$, см. Приложение 4), хотя и не такой большой как в 2002 году. Доля голосов СДПГ увеличилась в пострадавших регионах в среднем примерно на 2 процентных пункта на выборах 2005 года (в 2002 году было 7 процентных пунктов (Бехтеля и Хайнмюллера 2011)). Кроме того, при использовании метода *within*, а также CRSE, результаты в целом не меняются (см. Приложение 5-6).

При добавлении ковариат в регрессию эффект от наводнения также продолжает оставаться значимым, хоть и с небольшим понижением – в среднем увеличивает долю голосов СДПГ на 1,7 процентных пункта в пострадавших регионах (см. Приложение 7).

Однако на Рисунке 2 мы видим, что по сравнению с выборами в 1998 году доля голосов за партию СДПГ заметно снизилась во всех округах –, которые пострадали в 2002 году от наводнения ($Flooded = 1$) и которые не пострадали ($Flooded = 0$). Хотя разрыв между двумя группами сократился и возможно засчет помощи во время наводнения.

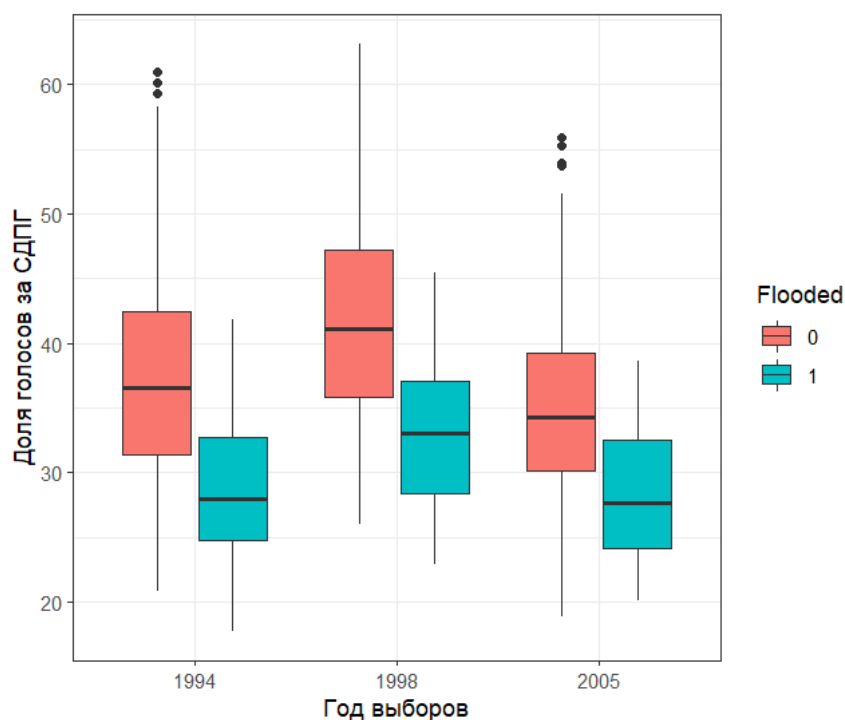


Рисунок 2. Диаграмма размаха доли голосов за СДПГ на выборах в 1994 году, 1998 году, 2005 году по двум группам округов: пострадавших во время наводнения 2002 года ($Flooded = 1$) и не пострадавших ($Flooded = 0$).

Вывод

Подводя итоги стоит отметить, что допущение о параллельности кажется вполне правдоподобным. Если говорить о непосредственном долгосрочном эффекте на поддержку СДПГ, то он сохраняется – увеличивает долю голосов за СДПГ в среднем примерно на 2 процентных пункта в пострадавших регионах.

Приложение

Проверка допущения

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	34.589	1.137	30.422	0.000
Flooded1	0.000	0.401	-0.001	0.999
PostPeriod1	4.610	0.133	34.704	0.000

Приложение 1. Линейная регрессия с дамми-переменными

	Estimate	Std. Error	t-value	Pr(> t)
Flooded1	0.00	0.401	-0.001	0.999
PostPeriod1	4.61	0.133	34.704	0.000

Приложение 2. Панельная регрессия с фиксированными эффектами (within)

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
Flooded1	-0.00040711	0.34292461	-0.0012	0.9991
PostPeriod1	4.61036210	0.13488000	34.1812	<2e-16 ***

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.' 0.1 ' ' 1

Приложение 3. Оценка с кластеризованными устойчивыми стандартными ошибками (CRSE)

Оценка эффекта

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	39.613	1.259	31.458	0
Flooded1	1.987	0.491	4.044	0
PostPeriod1	-6.773	0.153	-44.272	0

Приложение 4. Обычная регрессия с фиктивными переменными для электоральных округов

	Estimate	Std. Error	t-value	Pr(> t)
Flooded1	1.987	0.491	4.044	0
PostPeriod1	-6.773	0.153	-44.272	0

Приложение 5. Панельная регрессия с фиксированными эффектами (within)

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
Flooded1	1.98655	0.47297	4.2001	3.531e-05 ***
PostPeriod1	-6.77308	0.15309	-44.2431	< 2.2e-16 ***

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.' 0.1 ' ' 1

Приложение 6. Оценка с кластеризованными устойчивыми стандартными ошибками (CRSE)

t test of coefficients:

	Estimate	Std. Error	t value	Pr(> t)
Flooded1	1.791	0.414	4.325	<2e-16 ***
PostPeriod1	-6.504	0.218	-29.772	<2e-16 ***
popdens	1.460	1.095	1.334	0.183
SPD_inc	0.060	0.220	0.275	0.784
unemp_rate	0.114	0.122	0.933	0.351
net.migration	0.059	0.017	3.476	0.001 ***

Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.' 0.1 ' ' 1

Приложение 7. Оценка с кластеризованными устойчивыми стандартными ошибками (CRSE) с добавлением ковариат