

projet stats app séance du 04/04/25

3 avril 2025

1 Introduction

La base de données historique des inondations n'étant plus disponible, j'ai essayé d'évaluer la décote liée au risque d'inondation en utilisant la variation exogène de la zone inondable entre 2013 et 2020. Il y a en effet un tracé des zones inondables en 2013 et un autre en 2020. J'ai donc testé une estimation par double différence entre les logements passés d'un statut de non-inondable en 2013 à inondable en 2020.

Les deux années retenues pour la double différence sont 2019 et 2021. Ce qui restreint le nombre de transaction de 100000 à environ 35000.

Dans un premier temps, j'ai considéré comme groupe de traitement, les transactions passant de non-inondable pour tous les risque en 2019 à inondables en risque fort uniquement en 2021. Le groupe de contrôle étant quant à lui constitué des transactions restant non inondables.

Ci-dessous, le nombre de transactions suivant la zone de risque et l'année :

- risque débordement fort 2020 : 568
- risque débordement moyen 2020 : 1954
- risque débordement faible 2020 : 2814
- risque débordement fort 2013 : 545.0
- risque débordement moyen 2013 : 1093.0
- risque débordement faible 2013 : 1188.0

Il y a 113 transactions dans le groupe de traitement contre environ 11000 dans le groupe de contrôle.

TODO : cartographier le groupe de contrôle par rapport au groupe de traitement.

Comme le nombre de transactions est assez faible dans le groupe de contrôle, j'ai également considéré le passage de non-inondable à zone inondable à risque faible. Cette fois, le groupe de contrôle est composé de 811 transactions.

J'ai pas ailleurs restreint l'analyse au risque de débordement, car c'est pour ce risque que le plus de transactions sont concernées.

2 Modèle

Nous utilisons la régression linéaire suivante :

$$\log(\text{prix}/m_{it}^2) = \beta_0 + \beta_1 \times \text{Traitement} + \beta_2 \times \text{Post} + \beta_3 \times \text{Post} \times \text{Traitement} + \text{contrôles} + \varepsilon \quad (1)$$

Où Traitement représente une variable binaire pour l'appartenance au groupe de traitement. Post est aussi une variable binaire valant 1 pour l'année 2021 et 0 en 2019.

Les contrôles sont les suivants :

- distance à la mairie
- nombre de dépendances (variables indicatrices de 1 à 3)
- nombre de pièces principales (variables indicatrices de 1 à 6)
- surface batie
- surface terrain
- prix moyen dans la ville
- annee
- annee²

3 Résultats

Ci-dessous, les résultats pour le groupe de traitement composé des transactions passant de non-inondable à risque fort.

OLS Regression Results						
Dep. Variable:	log_prix_par_metre_carre	R-squared:	0.455			
Model:	OLS	Adj. R-squared:	0.455			
Method:	Least Squares	F-statistic:	566.8			
Date:	Thu, 03 Apr 2025	Prob (F-statistic):	0.00			
Time:	23:37:14	Log-Likelihood:	-5535.6			
No. Observations:	10861	AIC:	1.111e+04			
Df Residuals:	10844	BIC:	1.123e+04			
Df Model:	16					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	8.45e-06	4.08e-06	2.069	0.039	4.45e-07	1.65e-05
Traitement	0.0209	0.041	0.513	0.608	-0.059	0.101
Traitement_Post	0.1323	0.116	1.144	0.252	-0.094	0.359
Post	-0.0043	0.002	-2.069	0.039	-0.008	-0.001
distance_mairie_km	0.0150	0.003	5.743	0.000	0.010	0.020
surface_terrain	-3.816e-06	9.81e-07	-3.889	0.000	-5.74e-06	-1.89e-06
moyenne_prix_m2_ville	0.0005	6.04e-06	82.228	0.000	0.000	0.001
surface_reelle_bati	-0.0008	0.000	-5.901	0.000	-0.001	-0.000
annee	0.0085	0.004	2.069	0.039	0.000	0.017
annee_2	-2.642e-06	2.04e-06	-1.294	0.196	-6.64e-06	1.36e-06

dependance_1	0.1099	0.015	7.202	0.000	0.080	0
dependance_2	0.2219	0.023	9.684	0.000	0.177	0
dependance_3	0.3021	0.065	4.643	0.000	0.175	0
piece_principale_1	-0.0630	0.038	-1.668	0.095	-0.137	0
piece_principale_2	-0.0762	0.031	-2.490	0.013	-0.136	-0
piece_principale_3	0.0184	0.027	0.686	0.493	-0.034	0
piece_principale_4	0.1129	0.025	4.599	0.000	0.065	0
piece_principale_5	0.0787	0.024	3.307	0.001	0.032	0
piece_principale_6	0.0507	0.025	1.994	0.046	0.001	0

```

=====
Omnibus:                1090.148    Durbin-Watson:                1.881
Prob(Omnibus):           0.000    Jarque-Bera (JB):            1617.601
Skew:                    -0.766    Prob(JB):                     0.00
Kurtosis:                4.108    Cond. No.                    8.57e+23
=====

```

Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The smallest eigenvalue is 2.46e-31. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

Ci-dessous, les résultats pour le groupe de traitement composé des transactions passant de non-inondable à risque faible.

OLS Regression Results

```

=====
Dep. Variable:    log_prix_par_metre_carre    R-squared:                0.452
Model:            OLS                        Adj. R-squared:           0.451
Method:           Least Squares              F-statistic:              594.4
Date:             Thu, 03 Apr 2025            Prob (F-statistic):       0.00
Time:             23:37:14                   Log-Likelihood:           -5938.7
No. Observations: 11559                     AIC:                     1.191e+04
Df Residuals:     11542                     BIC:                     1.204e+04
Df Model:         16
Covariance Type:  nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975

Intercept	8.585e-06	4.08e-06	2.107	0.035	5.96e-07	1.66e-05
Traitement	-0.0089	0.018	-0.509	0.611	-0.043	0.025
Traitement_Post	0.0199	0.032	0.614	0.539	-0.044	0.084
Post	-0.0043	0.002	-2.107	0.035	-0.008	0.000
distance_mairie_km	0.0157	0.003	6.204	0.000	0.011	0.020
surface_terrain	-3.877e-06	9.39e-07	-4.131	0.000	-5.72e-06	-2.04e-06
moyenne_prix_m2_ville	0.0005	5.91e-06	83.713	0.000	0.000	0.001
surface_reelle_bati	-0.0008	0.000	-6.500	0.000	-0.001	0.000
annee	0.0087	0.004	2.107	0.035	0.001	0.016
annee_2	-2.709e-06	2.04e-06	-1.330	0.184	-6.7e-06	1.28e-06

dependance_1	0.1125	0.015	7.616	0.000	0.084	0
dependance_2	0.2270	0.022	10.125	0.000	0.183	0
dependance_3	0.2428	0.062	3.899	0.000	0.121	0
piece_principale_1	-0.0762	0.037	-2.072	0.038	-0.148	-0
piece_principale_2	-0.0753	0.030	-2.536	0.011	-0.133	-0
piece_principale_3	0.0232	0.026	0.893	0.372	-0.028	0
piece_principale_4	0.1187	0.024	4.996	0.000	0.072	0
piece_principale_5	0.0868	0.023	3.764	0.000	0.042	0
piece_principale_6	0.0611	0.025	2.469	0.014	0.013	0

=====						
Omnibus:	1147.095	Durbin-Watson:		1.897		
Prob(Omnibus):	0.000	Jarque-Bera (JB):		1695.558		
Skew:	-0.761	Prob(JB):		0.00		
Kurtosis:	4.097	Cond. No.		5.06e+23		
=====						

Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The smallest eigenvalue is 7.52e-31. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

4 limites

- coefficient causal non significatif
- pas de données antérieures à 2019 pour vérifier l'hypothèse de trend linéaire commun
- faible nombre de transactions dans le groupe de traitement