Appendix: Dismantling the "Jungle": Migrant Relocation and Extreme Voting in France

Paul Vertier*

Max Viskanic†

Matteo Gamalerio‡

This version: February 16, 2022

Abstract

Large migrant inflows have spurred anti-immigrant sentiment, but can small inflows have a different impact? We exploit the redistribution of migrants after the dismantling of the "Calais Jungle" in France to study the impact of the exposure to few migrants, which we estimate using difference-in-differences and instrumental variables. We find that in the presence of a migrant center (CAO), the growth rate of vote shares for the main far-right party (Front National (FN), our proxy for anti-immigrant sentiment) between 2012 and 2017 is reduced by about 12 percentage points. This effect, which crucially depends on the inflow's size, points towards the contact hypothesis (Allport 1954).

Keywords: Political Economy; Voting; Migration; EU; France; Migrants

JEL Classifications: C36, D72, J15, P16, R23

*LIEPP, paul.vertier@sciencespo.fr.

†LIEPP and IC Migrations, max.viskanic@sciencespo.fr.

‡Institut d'Economia de Barcelona (IEB), University of Barcelona, m.gamalerio@ub.edu.

[0]This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir program LIEPP (ANR-11-LABX-0091, ANR-11-IDEX-0005-02). We thank Yann Algan, Diane Bolet, Bjoern Brey, Julia Cagé, Guillaume Chapelle, Anthony Edo, Jean-Benoît Eyméoud, Roberto Galbiati, Ian Gordon, Sergei Guriev, Emeric Henry, Rachel Kranton, Ilyana Kuziemko, Mario Luca, Thomas Piketty, Panu Poutvaara, Paul Seabright, David Strömberg, Simon Weber, Ekaterina Zhuravskaya as well as participants at the Sciences Po Lunch Seminar, the CEPII Conference on Immigration in OECD countries, the 11th Cesifo Workshop in Political Economy,the Graduate Conference on Populism at the LSE, the lunch seminar in the PSPE Group at the LSE, the 3rd Workshop in Political Economy at the University of Bolzano, the Applied Lunch Seminar at PSE and the IEB Workshop on Political Economy for useful comments.

A1 Appendix: Data Description

Presidential election results for 2007, 2012, and 2017 at the municipality level come from the Ministry of Interior. We also use a dataset from Trendeo - Observatoire de l'investissement et de l'emploi (2017), which reports job destructions and creations at the municipal level in France between January 2009 and June 2017. This dataset provides a measure of local employment dynamics at the municipal level with high frequency. We use these variables as dependent variables in the empirical analysis.

We collected municipalities' characteristics from 2006, 2011, and 2013 French Censuses. We collected data on the total population, the share of individuals aged between 0 and 14 and over 60, and the share of individuals belonging to each of the eight official socio-professional categories (farmers, independent, white collars, intermediary professions, employees, blue collars, retired and inactive). Similarly, we consider the share of unemployment among the population aged between 15 and 64. Besides, we collected data on migrants' share of the total population, where migrants are defined as foreign-born individuals. We also collected the median disposable income by consumption unit (available only for municipalities of more than 50 inhabitants). To control for mayors' characteristics, we use the *Repertoire National des Elus*. This dataset provides information on the mayor's occupation, i.e., if she is a private employee or a civil servant, a teacher, a farmer, or an individual working in an industrial or liberal occupation. It also indicates the mayor's age, gender, and political orientation (e.g., whether the mayor is right-wing or not). We use these variables observed in the three years available in the data as time-varying covariates in the difference-in-differences analysis.[1]

We collected the number of hotel rooms in the municipality from INSEE. Data on municipalities located on the coast comes from the webpage Comersis. To control for the compliance of French mayors in implementing the CAOs, we use a list of mayors who declared to be willing to welcome migrants. This dataset, taken from the National French Television (France Télévision 2015), is neither official nor exhaustive but contains

[1]In practice, for every electoral year, we control for municipal and mayoral characteristics taken from the latest Census year that we were able to collect.

367 municipalities. From the *CIMADE*, we also collected information on the presence of other types of migrant centers, including CADA, HUDA, AT-SA, CPH, and PRAHDA. The data is most detailed for the CADA, for which we observe the number of places between 2012 and 2016. We computed the evolution of the number of places in CADAs at the municipality level with this information. Combining all this information with a GIS dataset of French municipalities (provided by the French National Geographic Institute (IGN)), we computed each municipality's distance to each of these centers.[2] We use the values of these variables in 2013 and, when available, their change between 2006 and 2013 as controls in our IV regressions to capture municipalities' current socio-economic and political conditions and their evolution after the 2008 financial crisis.

The location, number, and size of holiday villages are taken from the 2014 survey on tourism carried out by the French national statistical institute (INSEE). We collected data on group accommodation buildings such as homes for the elderly, disabled, drug addicts, and orphans from the National archive of health and social establishments (FINESS) for the year 2014. We use this information to build our instrument, which, importantly for the credibility of the instrument, is based on data observed before the dismantling of the Calais Jungle. Finally, we keep the municipalities without missing information in all these control variables in the empirical analysis.[3]

A2 Appendix: Main assumptions of Empirical Specifications

A2.1 Appendix: Difference-in-differences Approach

The main assumption of the difference-in-differences approach is that municipalities with and without a CAO should have been following the same electoral trends in the pre-

[2]Furthermore, we also computed the distance to the closest CAO for each municipality. We use this information in the difference-in-differences analysis to provide evidence about spillover effects in neighboring municipalities.

[3]The results do not qualitatively change if we keep these observations by replacing the missing values with the average value of the variable in the sample. Results can be made available upon request.

treatment period. We test this assumption in the Appendix A3, using data from the 2007 and 2012 presidential elections and the 2014 European and 2015 regional elections. Finally, an important factor behind the idea of this difference-in-differences model is that the *préfets* took the final decision about the location of CAO and not local politicians. Hence, it may be that the *préfets* decided the location of CAOs without being influenced by electoral trends. However, we cannot completely exclude the possibility that some mayors participated in the allocation process, generating doubts about the exogeneity of our treatment. For example, municipalities that volunteered to receive migrants and those with historically lower FN votes were more likely to host a CAO. We deal with this possibility and the fact that the assignment of the CAOs was not random by repeating the analysis using the instrumental variable approach described in the paper and in the next section.

A2.2 Appendix: Instrumental Variable Approach

OLS and diff-in-diff models may underestimate the effect of CAOs for two main reasons. First, as described in section 3.1 in the paper, we were not able to recover information on all the existing CAOs. This misinformation represents a measurement error that is likely to lead to an attenuation bias toward zero. Second, many municipal governments likely opposed the opening of CAOs for electoral reasons. For example, mayors from centrist and moderate parties may have done it to attract (or not lose) the votes of extreme and anti-immigrant voters. This potential movement of voters could lead to a negative correlation between CAOs and vote shares of mainstream parties. Thus, since we do not observe the bargaining process between municipalities and the government, simple OLS estimates may be biased towards zero. As described in the paper, to circumvent these potential biases, we propose to instrument CAOs location with the presence at the municipal level of pre-existing (i.e., built before the dismantling of the Calais Jungle) buildings that can accomodate groups of individuals.

The IV approach used in the paper relies on two main assumptions. The first assumption is that we need a first-stage regression in which $GroupBuildings_i$ correlates with the

presence of CAOs at the municipal level. In the empirical analysis, we formally test for this first assumption. The second assumption is the exclusion restriction one, which we can think of as being composed of two parts. The first part states that conditional on covariates, the instrument can be seen as good as randomly assigned (i.e., conditional on covariates, the instrument is independent of potential outcomes and treatment assignments). The second part requires that the instrument affects the outcome variable ΔFN only through the treatment CAO_i.[4] Relative to the first part, various concerns arise. For example, group accommodation buildings may have been built to accommodate the migrants redistributed after the dismantling of the Calais Jungle. For this reason, we measure the number of group accommodation buildings and holiday villages in 2014, before the beginning of the dismantling of the Calais Jungle. In addition, most of these buildings were built in the past and certainly not to host migrants.

An additional concern is that municipalities with group accommodation buildings and holiday villages may have different characteristics compared to municipalities without these buildings. For example, municipalities with group accommodation buildings may have a larger population or higher income per capita or may elect mayors with different political orientations. These differential characteristics may lead these municipalities to follow different electoral trends and experience a different probability of hosting migrants. For this reason, in the analysis, we control for a rich set of various municipal socio-economic and political attributes and trends that are likely to correlate with electoral trends and with the probability of hosting migrants. In addition, for what concerns holiday villages, we control for proxies for overall tourism (i.e., the log of the number of rooms in hotels and a dummy variable equal to 1 for municipalities on the coast), which is a factor that could correlate with electoral trends.

Relative to the second part, we show that our instrument correlates with the change in FN vote shares only between the 2017 and 2012 presidential elections, not between the 2012 and 2007 elections (see Table A1 in Appendix A3). This evidence suggests that

[4]For a discussion on how we can think of the exclusion restriction as being composed of two parts, see Angrist and Pischke 2009, chapter 4, page 117. Also, for a discussion of IV identification in constant-effect vs. heterogeneous-effect models and a detailed description of the independence and exclusion restriction assumptions, see Angrist and Pischke 2009, chapter 4, pages 150-153.

our instrument started to correlate with changes in electoral outcomes only when the buildings considered could be effectively be used to host migrants and not before that. In addition, we run a falsification test using Corsica's case: while this region has several holiday villages, it did not receive any CAOs. We do not find that municipalities in Corsica with a holiday village and group accommodation buildings had different voting trends for the FN between 2012 and 2017. Thus, while we cannot treat all this evidence as a formal test, these results suggest that the instrument started to correlate with the change in FN votes shares only during the dismantling of the Calais Jungle.

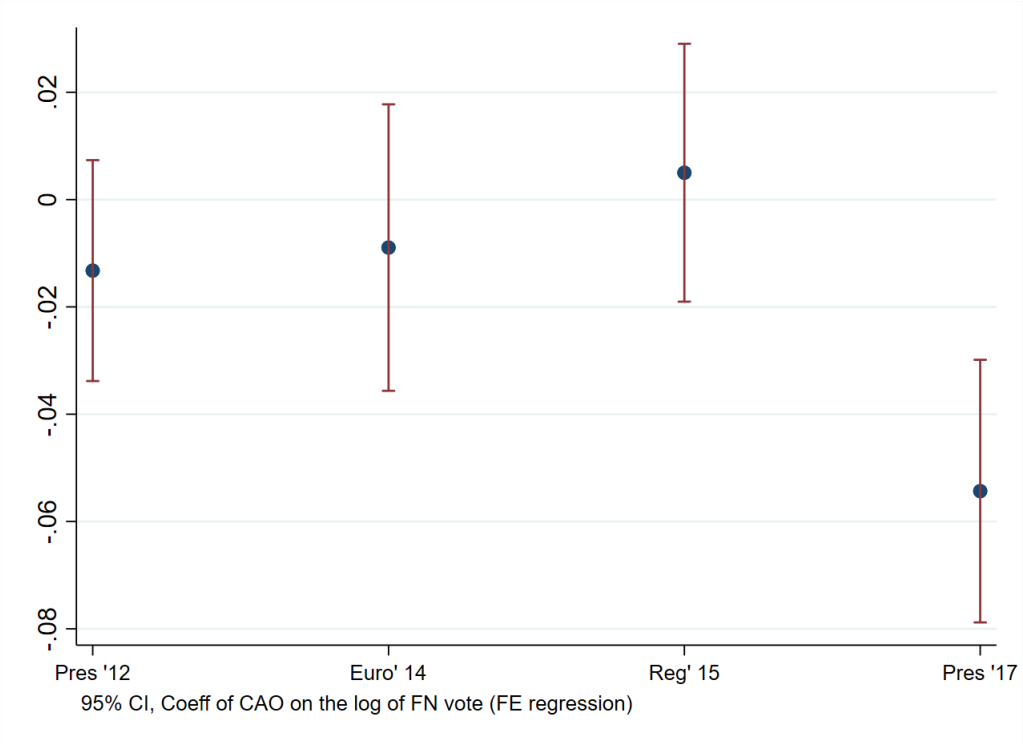
We argue that, since we control for a rich battery of municipal and mayoral characteristics, conditional on covariates, the instrument can be considered as good as randomly assigned. In addition, the evidence that the instrument correlates with the change in FN vote shares only between the 2017 and 2012 presidential elections suggests that it is plausible to think that group accommodation buildings and holiday villages affected the outcome variable ΔFN only through the treatment CAO_i. In conclusion, we think that the exclusion restriction assumption appears to be plausible in the context studied in this paper.

A3 Appendix: falsification and robustness checks

First, we consider whether we might be picking up pre-treatment electoral trends. We do this in two ways. First, we run a panel regression with municipal and year of election FE, where we evaluate the effect of CAO presence on various elections since 2012 (i.e., the Presidential elections of 2012, the European elections of 2014, the Regional elections of 2015, and the Presidential election of 2017). In Figure A1, where the effect of CAO in the Presidential elections of 2007 is normalized to be zero, the coefficient on CAO is never statistically different from zero except for the 2017 Presidential elections. This evidence shows that treated municipalities were not on different political trends before the election.

Second, in column 1 of Table A1, we run our diff-in-diff model using the 2007 and 2012 presidential elections, and we use the interaction between CAOs and the dummy for

Figure A1: Absence of Pretrends



the 2012 election as the main treatment. We do not find any effect. In column 2, we show that our instrument does not correlate (p-value equal to 0.950) with the log of the change in FN vote shares between the 2007 and 2012 elections. Finally, in column 3, we repeat the IV analysis using the log of the change in the FN vote shares as a dependent variable between the 2007 and 2012 elections. Also, in this case, we do not find any effect.

In Table A2, we consider Corsica, which represents an indirect test of our exclusion restriction. No migrants were relocated to Corsica, although it contains many holiday villages. Here, we regress our instrument on the change in voting outcomes for the FN in the French Presidential elections. Table A2 shows that no coefficient is significant. In addition, we report the p-value of the hypothesis test on the equality between the coefficient in front of the instrument in the reduced-form regression in the main sample (i.e., Table 1, Panel B, column 1) and the coefficient in Table A2, column 2. The p-value confirms that we reject the null hypothesis of equality between the two coefficients. Even tough this placebo test is run with a smaller number of obsevations compared to the main

Table A1: Pre-Trends: CAO Coefficients on Past Presidential Elections

	(1)	(2)	(3)
	Log(FG)	$\Delta FN_{2007-2012}$	$\Delta FN_{2007-2012}$
CAO x 2012	0.000		
	(0.013)		
Group buildings		0.000	
		(0.000)	
CAO			0.003
			(0.040)
Model	DiD	Reduced form	IV
Covariates	Yes	Yes	Yes
Time FE	Yes	No	No
Municipal FE	Yes	No	No
Regions FE	No	Yes	Yes
F-statistic	-	-	41.05
Observations	56,195	26,884	26,884

Difference-in-differences estimates in column 1, Reduced form model in column 2, Instrumental variables estimates in columns 3. Variables reported in the Table: CAO = 1 for a migrant center in the municipality; 2012 = 1 for 2012 presidential election; Group buildings = number of village vacances and group accommodation buildings in the municipality. Control variables in column 1: municipality sociodemographic characteristics, the mayor's party, and personal characteristics. Control variables in columns 2-3: municipality sociodemographic characteristics (in 2013 and evolution between 2006 and 2013), coastal dummy variable, the log of the number of hotel rooms, whether the municipality volunteered to receive migrants, the log of distance to the closest permanent migrant center, the evolution of the number of places in CADAs, the mayor's party and personal characteristics. Standard errors clustered at the municipality level in parentheses in column 1. Standard errors clustered at the *département* level in parentheses in columns 2-3. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

sample (i.e., 200 vs. 26,888 observations), we think that these additional results underline the validity of our IV approach.

Table A2: No Link between Holiday Villages and FN trend in Corsica

	(1) Δ_{FN}	(2) Δ_{FN}
Group buildings	0.000	0.002
	(0.000)	(0.002)
P-value difference		0.026
Regression	OLS	OLS
Controls	No	Yes
Observations	352	200

Columns 1 to 2 report the results of OLS regressions of the variation of log FN votes between the presidential elections of 2012 and 2017 on the dummy for a holiday village and group accomodation buildings. Control variables in columns 2-3: municipality sociodemographic characteristics (in 2013 and evolution between 2006 and 2013), coastal dummy variable, the log of the number of hotel rooms. Standard errors clustered at the *département* level in parentheses.The p-value reported in the Table is the p-value of the hypothesis test of the equality between the coefficent in front of Group buildings in Table 1, Panel B, column 1 and the coefficent in this Table, column 2. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In Table A3, we rule out the possibility that changes in electoral turnout explain the effect of CAOs on FN and FG vote shares. Specifically, we split the sample between municipalities that experienced a negative change in electoral turnout and those that experienced a positive change. As we can see, the coefficients for both FN and FG vote shares are similar in magnitude for the two groups of municipalities. Besides, the coefficients are statistically different from zero only for municipalities that experienced a drop in electoral turnout. This evidence suggests that changes in electoral turnout and voters mobilization do not appear to be the main driver of our results.

Table A3: The role of Turnout and mobilization

	(1)	(2)	(3)	(4)
	Δ_{FN}	Δ_{FN}	Δ_{FG}	Δ_{FG}
Sample	$\Delta_{Turnout} < 0$	$\Delta_{Turnout} > 0$	$\Delta_{Turnout} < 0$	$\Delta_{Turnout} > 0$
CAO	-0.126***	-0.112	0.159***	0.167
	(0.038)	(0.085)	(0.060)	(0.109)
Observations	19,366	7,522	19,359	7,519
F-statistic	48.92	6.249	48.91	6.249
Model	IV	IV	IV	IV
Covariates	Yes	Yes	Yes	Yes
Regions FE	Yes	Yes	Yes	Yes

Instrumental variables estimates in all columns. Variables reported in the Table: CAO = 1 for a migrant center in the municipality. Control variables in all columns: municipality sociodemographic characteristics (in 2013 and evolution between 2006 and 2013), coastal dummy variable, the log of the number of hotel rooms, whether the municipality volunteered to receive migrants, the log of distance to the closest permanent migrant center, the evolution of the number of places in CADAs, the mayor's party and personal characteristics. Standard errors clustered at the *département* level in parentheses in all columns. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In Table A4, we check that few outliers do not drive the first-stage relationship. As reported in the bottom panel of Table A4, the instrument takes values that go from 0 up to 274. Hence, one potential concern is that a few municipalities with big values of the instrument drive the first-stage relationship. To rule out this possibility, following the example of Steinmayr (2020), we rerun the first-stage regression dropping respectively the municipalities/observations with the 5, 10, and 30 biggest values of the instrument (i.e., municipalities with respectively more than 191, 132, and 85 group accommodation buildings). As we can see from columns 1-3 of Table A4, dropping these observations, we get first-stage regressions with even bigger F-statistics. In addition, in column 4 of Table A4, we have rerun the first stage regression using the instrument winsorized at the 99th percentile (i.e., considering the sample of municipalities with at least one group accommodation building, the 99th percentile corresponds to 43 group accommodation buildings). We get a first-stage relationship with a bigger F-statistic even in this case. This evidence suggests that few outliers do not drive the first-stage relationship.

Finally, in Figure A2, we use the methodology developed by Hainmueller et al. (2019)

Table A4: The role of outliers in the first-stage regression

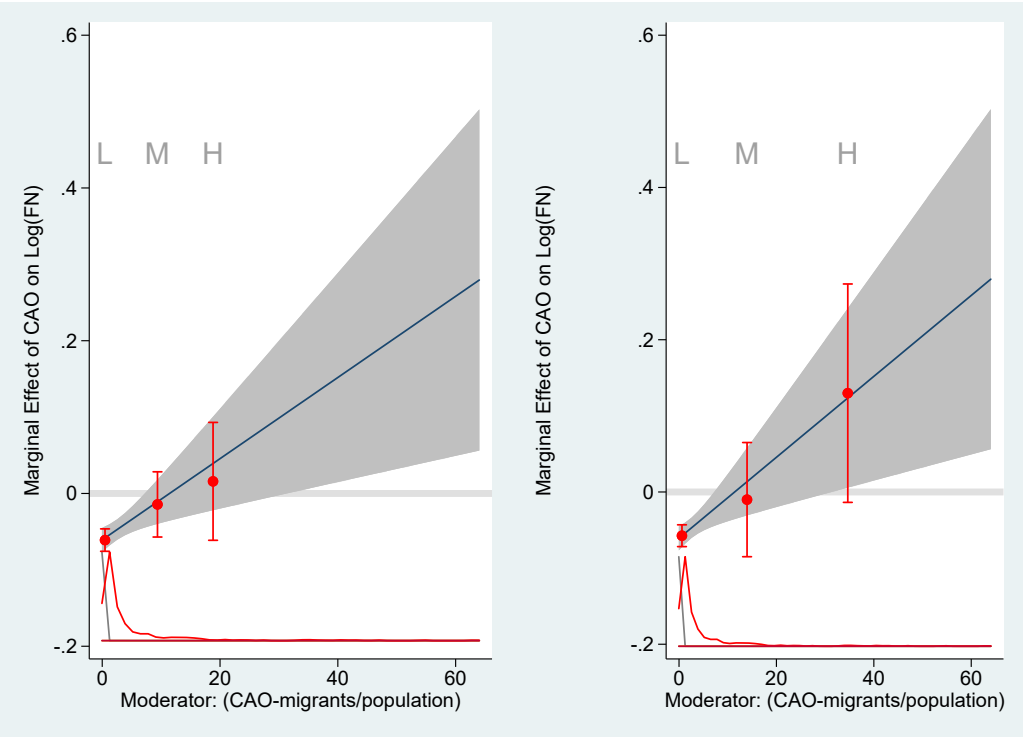
	(1)	(2)	(3)	(4)
Dependent variable	CAO	CAO	CAO	CAO
Covariates	Yes	Yes	Yes	Yes
Regions FE	Yes	Yes	Yes	Yes
Sample	Drop 5	Drop 10	Drop 30	Winsorize
Group buildings	0.006***	0.007***	0.007***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	26,883	26,877	26,858	26,888
F-statistic	86.25	85.30	42.15	56.56
Distribution of Group buildings				
	Mean	Standard deviation	Min	Max
Group buildings	1.16	5.96	0	274
Group buildings if Group buildings > 0	4.09	10.65	1	274

First-stage regressions in all columns. Variables reported in the Table: CAO = 1 for a migrant center in the municipality; Group buildings = number of village vacances and group accommodation buildings in the municipality. Control variables in all columns: municipality sociodemographic characteristics (in 2013 and evolution between 2006 and 2013), coastal dummy variable, the log of the number of hotel rooms, whether the municipality volunteered to receive migrants, the log of distance to the closest permanent migrant center, the evolution of the number of places in CADAs, the mayor's party and personal characteristics. Samples used: 1) Drop 5 = we drop the 5 municipalities/observations with the highest values in Group buildings; 2) Drop 10 = we drop the 10 municipalities/observations with the highest values in Group buildings; 3) Drop 30 = we drop the 30 municipalities/observations with the highest values in Group buildings; 4) Winsorize = we winsorize Group buildings at the 99th percentile. Standard errors clustered at the *département* level in parentheses in all columns. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

to test for the plausibility of a linear interaction effect in the heterogeneity analysis described in Table 3. Figure A2 shows the relationship between the marginal effect of CAOs on FN vote shares and the moderator, which is the number of migrants in CAOs every 1000 inhabitants, standardized as to take mean 0 and standard deviation 1. We implement this analysis applying the Stata command interflex to model 1. This test compares the conditional effect estimates from a binning estimator with those from a standard multiplicative interaction model. We use different cutoffs to split the sample of municipalities with a CAO in the three bins required by the binning estimator. First, in the graph on the right, we split the sample between municipalities with a size of the inflow below 7 (i.e., 27 migrants per 1000 inhabitants) standard deviations, those between 7 and 12 (i.e., 47 migrants per 1000 inhabitants) standard deviations, and those above 12 standard deviations. We pick these two thresholds because they represent the cutoffs at which, according to the estimates in Table 3, the effect of CAOs on FN votes becomes first insignificant and then switches sign. Second, in the left graph, we use as thresholds 12 and 17 (i.e., 68 migrants per 1000 inhabitants) standard deviations. These thresholds represent the 90th and the 95th percentile of the distribution of the inflow size, respectively, if we consider only municipalities with a CAO. We use these thresholds so that to split in samples of a similar size those municipalities above the threshold at which, according to our estimates, the effect of CAOs on FN votes switches sign.[5] As we can see, the estimates from the binning estimator appear to sit on the estimated linear marginal effect line in both graphs. This evidence reinforces the plausibility of a linear interaction effect.

[5]We have also tried to use other combinations of thresholds, and we get similar results. For example, we have tried with 7 and 17 standard deviations, 7 and 59 (i.e., the threshold at which the effect becomes positive and statistically significant), and 12 and 59 standard deviations. Results can be made available upon request.

Figure A2: Robustness check interaction model



References

Gordon Allport. The nature of prejudice. *Addison-Wesley.*, 1954.

Joshua Angrist and Steve Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press, 2009.

France Télévision. Carte. quelles sont les communes volontaires pour accueillir des réfugiés ? by baïetto, t. and dalbosco, m. 2015.

Jens Hainmueller, Jonathan Mummolo, and Yiqing Xu. How much should we trust estimates from multiplicative interaction models? simple tools to improve empirical practice. *Political Analysis*, vol. 27:163–192, 2019.

Andreas Steinmayr. Contact versus exposure: Refugee presence and voting for the far-right. *Review of Economics and Statistics*, 2020.

Trendeo - Observatoire de l'investissement et de l'emploi. 2017.