

APPENDIX. THE LOGISTIC REGRESSION MODEL ON THE DECISION TO MAINTAIN IN THE FRENCH NATIONALITY BY THE COMMISSION FOR THE REVIEW OF NATURALIZATIONS

This model is an attempt to distinguish the role played by different criteria in the Commission's decision making. It was realized with Pierre Mercklé, to whom I would like to express my sincere thanks. Due to the low number of withdrawals during the first examination of files, I chose to model decisions to maintain, as opposed to decisions to launch an investigation or withdraw. Of course, the latter two are not interchangeable in terms of the consequences for individuals. But modelling one as opposed to the other would lead above all to an opposition over time since withdrawals decided at first instance become quite a minority from 1942 onward. It would therefore seem preferable to reason about the Commission's amenity, when it decided to maintain French nationality, versus its suspicion, whether it took the form of a request for additional information or a notice of withdrawal. The model is constructed on the basis of one unit per naturalization file; the variables therefore qualify the head of household and not all the individuals in the family, which corresponds, as we have seen, to the Commission's review practices.

I selected seven explanatory variables: country of origin, profession, matrimonial status, year of naturalization, the number of the specific subcommission that examined the file, the time of day of review (morning or afternoon), and the name of the judge. As we have seen, it is not possible to construct a stable variable that determined whether a given individual is “perceived as Jewish” or not. In fact, these designations resulted from a cluster of criteria—onomastic, professional, national—and they were only rarely mentioned in the reports.

Table A reports the results for a binomial logit model of decision type. The dichotomous model “Decision to maintain” vs. “Other decision” converges on the selected explanatory variables thus prepared. The estimate of the model is a sub-sample of 374 of the 495 files in the initial sample: 3 were removed due to missing values for the explained variable; 118 individuals were removed due to missing values for the explanatory variables.

The test of the nullity of all the modalities of the explanatory dimensions of this model indicates only two significant explanatory variables in this specific order: place of birth and name of the judge (table A). The model cannot rule out the effect of the other explanatory dimensions on the decision, all other things being equal. A variable is considered as “statistically significant” in a regression when the observed effect is probably not due to chance alone, given the size of the sample. The last column on the right side of the table uses asterisks to denote the low probability that the observed effects are due to chance alone. * means that this probability is less than 5 percent, **less than 1 percent, ***less than 0.1 percent. The absence of an asterisk means that *no statistically significant effect* can be measured for the corresponding category on the dependent variable. This means either that there is no effect or that our sample is too small to observe it.

We then look at the results to see which modalities of these variables appear significant in the regression model (table B). When there is an asterisk in the last column on the right of the table, hence a statistically significant effect, the sign of the coefficient in the column (1) of the table indicates the direction of the effect. Variables with positive coefficients increase the probability of being maintained in the French nationality.

Variables with negative coefficients decrease it. Large (whether positive or negative) coefficients signal the strongest effects. In order to interpret the coefficients of a logistic regression, it is usual to transform them into “odds ratios” (an odds ratio is calculated as the exponential of the coefficient). Positive coefficients produce odds ratios superior to 1, meaning that the chance of being maintained in the French nationality is x times more likely for the considered category, as compared with reference. Negative coefficients produce odds ratios between 0 and 1: in this case, the opposite of the odds ratio is calculated (in parentheses in Table B), meaning that the chance of not being maintained is X times less likely for the considered category, as compared with reference.

Note that it can be quite interesting to comment on “non-effects”: in this case, for instance, the profession, whereas one might have expected these factors to play a role. This does not mean that the percentage of maintain decisions was not higher for some professions than for others, as we have seen. But there is no measurable influence, *all other things being equal*, of the profession on the decision to maintain in the French nationality.

The reference variable is the one in relation to which the effects are calculated; chosen here are the most frequent ones. Thus, with regard to place of birth, those born in Asia Minor are 31 times less likely to obtain maintenance than those born in Italy (reference category), and those born in Eastern Europe are 16 times less likely to obtain it. On the other hand, those born in Southern Europe are 3 times more likely to obtain it (the latter coefficient is nevertheless significant at the 10% threshold only).

Table A. Test of the nullity of all the modalities of the explanatory dimensions

	Df	Significance
<none>		NA
Year of naturalization	5	
Country of birth	8	***
Profession	6	
Matrimonial status	1	
Number of subcommission	3	
Name of the judge	22	***
Time of day	1	

Note: The table uses asterisks to denote the low probability that the observed effects are due to chance alone. *** means that this probability is less than 0.1 percent. The absence of an asterisk means that *no statistically significant effect* can be measured for the corresponding category on the dependent variable. This means either that there is no effect or that our sample is too small to observe it. Boldface type indicates the significant modalities.

Table B. Determinants of the decision to maintain nationality

Explanatory modalities	Coefficient	p value	Odds ratio	Significance
YEAR OF NATURALIZATION				
<i>Before 1930</i>	<i>REF</i>			
1931-1932	-0,70	0,24	0,49 (2)	
1933-1935	+0,53	0,412	1,70	
1936	-14,34	0,997	0,00 (1696299)	
1937-1938	-1,01	0,433	0,36 (2,75)	
1939-1940	+1,21	0,298	3,34	
COUNTRY OF BIRTH				
Africa	+0,36	0,662	1,43	
America	+1,10	0,525	3	
Asia Minor	-3,43	0,002	0,03 (31)	**
Belgium	+1,18	0,228	3,25	
Northern Europe	-0,10	0,873	0,90 (1)	
Eastern Europe	-2,75	0	0,06 (16)	***
Southern Europe	+1,08	0,067	3	
France	+1,09	0,179	3	
<i>Italy</i>	<i>REF</i>			
PROFESSION				
Agrarian sector	-0.28	0.556	0.76 (1.32)	
Employee	+0.01	0.989	1.01	
Foreman	+12.61	0.997	298646.54	
<i>Laborer</i>	<i>REF</i>			
Intellectual profession	-1.77	0.054	0.17 (5.90)	
Independent profession	-0.23	0.557	0.79 (1.26)	
Other	+0.59	0.542	1.81	
MATRIMONIAL STATUS				
Single	-0,31	0,428	0,73 (1,5)	
<i>Married</i>	<i>REF</i>			
NUMBER OF SUBCOMMISSION				
<i>C1</i>	<i>REF</i>			
C2	-1,05	0,228	0,35 (3)	
C3	-0,19	0,861	0,83 (1,2)	
Plenary	-1,07	0,535	0,34 (3)	
NAME OF THE JUDGE				
ALBUCHER	+1,44	0,261	4,22	
<i>BERTHELEMOT</i>	<i>REF</i>			
CHERON	+1,11	0,427	3,05	
COMBIER	+0,77	0,618	2,15	
COUPILLAUD	-0,06	0,964	0,95 (1)	
DARRAS	+1,29	0,53	3,65	
DUMONCEAU	+0,35	0,802	1,42	
FLEURY	-0,20	0,924	0,82 (1)	
GERMAIN	+0,62	0,579	1,85	
LEGENDRE	+0,55	0,663	1,74	
MARTIN	-16,10	0,992	0,00 (9822298)	
MOUSSARD	-0,85	0,533	0,43 (2,5)	

PAGENEL	+19,83	0,996	409432480	
PAPON	+0,07	0,95	1,08	
PARLANGE	+0,67	0,72	1,95	
POTTIER	-2,17	0,127	0,11 (9)	
SENGENCE	+1,65	0,159	5,18	
SEYER	-17,13	0,994	0,00 (27507928)	
SIRE	+3,20	0,039	24,5	**
THIRION	+0,35	0,808	1,42	
VALLEE	+18,36	0,996	93844186,30	
VAURY	-15,80	0,993	0,00 (7259668,32)	
VIEILLEDENT	+2,95	0,04	19	**
VOULET	-0,99	0,526	0,37 (2,68)	
TIME OF DAY				
Morning	<i>REF</i>			
Evening	-0,21	0,803	0,81 (1,23)	

Tests	Value
N	374
AIC	399,70
Pseudo-R2	0,39
C	0,89
Somer's D	0,78

Note : The table uses asterisks to denote the low probability that the observed effects are due to chance alone. * means that this probability is less than 5 percent, **less than 1 percent, ***less than 0.1 percent. The absence of an asterisk means that *no statistically significant effect* can be measured for the corresponding category on the dependent variable. This means either that there is no effect or that our sample is too small to observe it. Boldface type indicates the significant modalities.