**PS : Political Science & Politics Symposium**

**Forecasting the 2022 French Presidential Election**

**with a SUR Regionalized Model**

PS-D-21-00244

Appendix

**1°) Classification of the political blocks used to build the dependent variables**

Table 1. Dependent variable for each equation in the SUR model (following Jérôme, Mongrain et Nadeau, 2022)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Election** | **Incumbent** **(INC)** | **Opposition** **(OPP)** | **Challenger** **(CHAL)** | **Diverse right** **(DVR)** | **Far right** **(FAR)** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| 1974 | C.-Delmas (CON) | Total left | Giscard (CEN)3 | DVR | J.-M. Le Pen |
| 1981 | Giscard (CEN) | Total left | Chirac (CON) | – | – |
| 1988 | Chirac (CON)1,6 | Total left | Barre (CEN)3 | – | J.-M. Le Pen |
| 1995 | Balladur (CEN)1,6 | Total left | Chirac (CON) | DVR | J.-M. Le Pen |
| 2002 | Jospin (PS)1,6 | CON + CEN | Non-PS left4 | DVR | J.-M. Le Pen |
| 2007 | Sarkozy (CON) | Total left | Bayrou (CEN)2 | DVR | J.-M. Le Pen |
| 2012 | Sarkozy (CON) | Total left | Bayrou (CEN)2 | DVR | Marine Le Pen |
| 2017 | Hamon (PS) | CON | Non-PS left + CEN5 | DVR | Marine Le Pen |
| 2022 | Macron (CEN) | Total left | Pécresse (CON) | DVR | FAR (M. Le Pen) |

*Note*. (1) Prime minister. (2) In opposition. (3) In majority. (4) Non-PS left = Far left + Communist Party (or Front de Gauche or France insoumise) + Ecologists. (5) In 2017, the Socialist Party had two challengers: the non-PS left and the center with François Hollande’s former minister of the Economy, Emmanuel Macron, allied with François Bayrou.

(6) According to the constitution, in case of cohabitation, the Prime Minister leads and conducts the economic policy (and homeland policy) of the Nation. Therefore, the new majority applies the program on which it was elected. The President only covers international affairs (of a low weight in the French vote) and simply convenes the Council of Ministers. In 1986 and 1995, the Socialist Party found itself in opposition to the government, as did the Right in 2002. The Prime minister is then accountable for its record, not the President.

**2°) Acronyms for geographical areas**

Auvergne-Rhône-Alpes (ARA)

Bourgogne-Franche-Comté (BFC);

Bretagne (BRE);

Centre-Val-de-Loire (CVL);

Grand Est (GES);

Hauts-de-France (HDF);

Ile-de-France (IDF);

Normandie (NOR);

Nouvelle-Aquitaine (NAQ);

Occitanie (OCC);

Pays de la Loire (PDL);

Provence-Alpes-Côte d’Azur (PAC).

Corse (COR);



IGN, France

**3°) Definition and list of cited political parties and names**

* Conservative parties (Gaullists)

UDR : Union des démocrates pour la République (1967-1976)

RPR : Rassemblement pour la République (1976-2002)

UMP : Union pour un Mouvement Populaire (2002-2015)

LR : Les Républicains (2015-….)

2022 Candidate : Valérie Pécresse

* Centrists

UDF : Union pour la Démocratie Française (1978-2007)

MoDEM : Mouvement Démocrate (2007-….)

2022 candidate : Emmanuel Macron

LREM : La République en marche ! (2017-….)

 2022 candidate : Emmanuel Macron

* Far Left, Left, diverse Left and Greens

LO : Lutte ouvrière (Far-Left) (1968-….)

2022 candidate : Nathalie Arthaud

LCR : Ligne communiste révolutionnaire (Far-Left) (1974-….)

2022 candidate : Philippe Poutou

FG : Front de Gauche (2009-2016) (Far-Left and Communists)

PC : Parti communiste (1943-….)

2022 candidate : Fabien Roussel

LFI : La France Insoumise (2016-….) (Far-Left and Communists)

2022 Candidate : Jean-Luc Mélenchon

PS : Parti socialiste (1969-….)

2022 candidate : Anne Hidalgo

Divers gauche : 2022 candidate : Arnaud Montebourg,

EELV : Europe Ecologie les Verts (2010-….)

2002 candidate : Yannick Jadot

* Diverse Right

Reconquête ! : (2021-….)

2022 candidate : Eric Zemmour

DLF : Debout la France (1999-….)

2022 candidate : Nicolas Dupont-Aignan

UPR : Union populaire Républicaine (2007-….)

2022 candidate : François Asselineau

Résistons ! (2016-….)

2022 candidate : Jean Lassalle

* Far-Right

Les Patriotes : (2017-….)

2022 candidate : Florian Philippot

FN : Front National (1972-2018)

RN : Rassemblement National (2018-….)

2002 candidate : Marine Le Pen

**4°) Econometric Methodology : the SUR regression**

The seemingly unrelated regression (SUR) approach (see Zellner 1962) is used to estimate the vote share of political blocs simultaneously in the first round and then the two qualified candidates for the second round. Seemingly unrelated regressions models have already been applied to British, German, and French legislative elections (Arzheimer and Evans 2010; Jérôme, Jérôme-Speziari, and Lewis-Beck 2017; Mongrain 2021) and offer the possibility to forecast the results of multiple parties. The SUR model therefore takes into consideration the possibility that the determinants of voting for one party may also help to determine the vote for other parties. If *distinct* OLS equations are proposed for each party, then it must be assumed that the error term of an equation (which roughly represents the relevant factors omitted from the analysis) is not correlated with the error term of the other equations. This is highly improbable when using election results data. The SUR model attempts to correct this problem by taking into account the correlations between the errors of the different equations. Since more than two parties are able to obtain respectable proportion of votes in the first round in French presidential elections, relying on the SUR methodology is a worthwhile option (see Timm 2016).

Arzheimer, Kai, and Jocelyn Evans. 2010. “Bread and Butter à la française: Multiparty Forecasts of the French Legislative Vote (1981–2007).” *International Journal of Forecasting* 26 (1): 19–31.

Jérôme, Bruno, Véronique Jérôme-Speziari, and Michael S. Lewis-Beck. 2017. “The Grand Coalition Reappointed but Angela Merkel on Borrowed Time.” *PS: Political Science and Politics* 50(3): 683–685.

Mongrain, Philippe. 2021. “10 Downing Street: Who’s Next? Seemingly Unrelated Regressions to Forecast UK Election Results.” *Journal of Elections, Public Opinion and Parties* 31 (1): 22–32.

Timm, Neil H. 2002. *Applied Multivariate Analysis*. New York: Springer.

Zellner, Arnold. 1962. “An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias.” *Journal of the American Statistical Association* 57 (298): 348–368.

**5°) List and definition of variables**

* **SUR Model (First round)**

*Dependant variables*

INC : vote percentage for the party supporting the outgoing executive in a region i. The incumbents are the conservatives in 1974, 1988, 2007 and 2012; the socialists[[1]](#endnote-1) in 2002 and 2017; and the centrists in 1981, 1995 and 2022 (for forecasting).

Jacques Chirac (Prime Minister) in 1988, Edouard Balladur (PM) in 1995 and Lionel Jospin (PM) in 2002 are considered as the real incumbents, as they were responsible for the majority in a situation of cohabitation. They are accountable for the executive’s record in the eyes of voters, unlike the President under cohabitation. It should be noted that Balladur was more accurately the "candidate of the centrists" because he was supported by the UDF against Chirac (RPR).

For 2022, we have followed two criteria to classify Macron in the center. First, since the 2020 municipal elections, La République en Marche is classified as "Diverse centers" by the Ministry of the Interior. Second, Macron's electorate is now composed of two thirds of center-right voters and one third of center-left (it was the opposite in 2017).

OPP : vote percentage for the “natural” opposition in a region i. The opposition is formed by the total left from 1974 to 1995, from 2007 to 2012 and then 2022 (forecast); the conservatives plus the center in 2002, then the conservatives in 2017.

CHAL : vote percentage for the challengers competing with the incumbents within the same camp in a region i, our third equation. The center in 1974, 1988, 2007 and 2012 will be considered as the challenger to the chief executive’s party; the conservatives in 1981, 1995 and 2022 (forecast); the non-socialist left in 2002 and the non-socialist left plus the center in 2017.

In 2017, the PS candidate Benoît Hamon was “challenged” by the center (Emmanuel Macron, outgoing ex-minister) and the non-socialist left (far left, Communist Party, Front de Gauche, France Insoumise, and greens).

FAR : vote percentage for the far-right in a region i. Jean-Marie Le Pen in 1974 then from 1988 to 2007 and Marine Le Pen from 2007 to 2022 (forecast).

DVR : vote percentage for the “diverse-right” (DVR) in a region i present in 1974 and from 1995 to 2022 (forecast). Éric Zemmour is included with the “diverse-right” because he locates himself at the intersection between the classical right and the nationalist right.

*Independant variables*

U6 : regional unemployment rate in the month of December preceding the election (Insee)

DU : variation in regional unemployment from December t-1 to December t-2 (Insee)

INCLG : vote percentage for the party supporting the outgoing executive in a region i at the past legislative elections (1973, 1978, 1986, 1993, 1997, 2002, 2007, 2012 and 2017 for the forecast).

OPPLG : vote percentage for the “natural” opposition at the past legislative elections (1973, 1978, 1986, 1993, 1997, 2002, 2007, 2012 and 2017 for the forecast).

CHALLG : vote percentage for the challengers competing with the incumbents within the same camp in a region i at the past legislative elections (1973, 1978, 1986, 1993, 1997, 2002, 2007, 2012 and 2017 for the forecast)

FARLG : vote percentage for the far-right in a region i at the past legislative elections (1973, 1978, 1986, 1993, 1997, 2002, 2007, 2012 and 2017 for the forecast).

INCSH : incumbent strongholds index (see joined database) at the first round. The electoral strongholds index is a sliding weighted dummy. Example: starting in 1974, when in a region a party or ideological camp obtained a score higher than the 1974 average, the index starts at 1. Then, if this score is below the 1981 average, the index falls to 0.5 (1/2). If this score rises above the average, the index rises to 0.66 (2/3). We continue in this way until 2017 (8th observation point). For the forecast, the 2017 value of the index is used for 2022.

OPPSH : « natural » opposition strongholds index (see joined database and above explanations) at the first round.

CHALSH : challengers (competing with the incumbents) strongholds index at the first round. Before building this index, electoral scores from different parties are sometimes aggregated to follow the classification described in 1°): in 2002 for non-PS Left and for 2017 considering Benoît Hamon has two challengers, namely the non-PS Left + Center (Macron).

FARSH : Far-Right strongholds index at the first round.

DVRSH : Diverse Right strongholds index at the first round.

INCPOP : popularity of the executive in the quarter before the election (Ifop JDD).

LEPNPOP : popularity) of the RN (or FN) candidate (Sofres-Kantar-One Point).

DVRPOLL : average vote intentions in the polls (%) in Diverse droites’ favor in the first quarter before the election

DUMHOLL : dummy variable scored 1 (0 otherwise) coding the impact of François Hollande’s defection in 2017

DUMROYER : dummy variable coded (0 otherwise) measuring the equivalent of an « homestate advantage » for Jean Royer in 1974 in his stronghold of Centre Val de Loire region.

DUMVILLIER : dummy variable coded (0 otherwise) measuring the equivalent of an « homestate advantage » for Philippe De Villiers in 1995 in his stronghold of Pays de la Loire region.

MODEM : dummy variable coded 1 (0 otherwise) to measure the electoral consequences when François Bayrou sacrificed the UDF to create the MoDEM (MODEM).

CONSOPP : dummy variable coded 1 (0 otherwise) to measure how the diverse right-wingers benefit from the presence of the conservatives in the opposition

PSINC : dummy variable coded 1 (0 otherwise) when the PS is incumbent (in 2002 and 2017)

COHAB : dummy variable coded 1 (0 otherwise) to measure the electoral cost of cohabitation for outgoing Prime Ministers

CORS.. : dummy variable coded 1 (0 otherwise) for Corsica in 1974, 1988, 2002, 2007, 2012 et 2017.

* Far-Right Model (second round)

*Dependant variable :*

RNT2 : vote percentage for the RN-FN in the second round in 2002 and 2017.

*Independant variables :*

POLEPEN : voting intentions for the leader of the far right in the second round against Chirac (2002) or Macron (2017)

RNSHT2 : regional strongholds of the RN-FN in the second round

CORS17 : dummy variable coded 1 (0 otherwise) for Corsica in 2017

* Classic right Model (second round)

*Dependant variable :*

RP2 : vote percentage for the conservative candidate in the second round in 1974-1995 et 2007-2012.

*Independant variables :*

RP1QUAL : vote percentage for the qualified conservative candidate in the first round

COMP1: vote percentage for the centrist candidate in the first round

DVR: vote percentage for the Diverse Right candidates in the first round

FAR: vote percentage for the Far-Right candidate in the first round

ZFDRO: strongholds index for the classic Right in the second round

**6°) Data used to forecast 2022 presidential election**



















**7°) Source of the data**

**Electoral data**

Presidential and legislative elections from 2002 to 2017

<https://www.data.gouv.fr/fr/posts/les-donnees-des-elections/>

Presidential elections 1974, 1981, 1988 and 1995

Legislative election 1973, 1978, 1981, 1986, 1988, 1993 and 1997

Source Le Monde : Electoral booklets published the day after the elections and compilation calculations by the authors.

**Unemployment data :**

From 1982 to 2021

<https://www.insee.fr/fr/statistiques/2012804>

From 1973 to 1982

Bulletin mensuel des statistiques du travail (INSEE)

Compilation work by the authors

**Strongholds data**

Construction of « areas of strength » index for each political family by the authors (see attached databases)

**Popularity data for President and Prime minister (1973-2021)**

Source : Ifop France soir then Ifop Journal du Dimanche for the President and the Prime minister

**Popularity data for Jean-Marie Le Pen and de Marine Le Pen**

Source : Sofres then Kantar Sofres then Kantar One Point

**Polling data on the French presidential election**

<https://fr.wikipedia.org/wiki/Liste_de_sondages_sur_l%27%C3%A9lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_2022>

<https://fr.wikipedia.org/wiki/Liste_de_sondages_sur_l%27%C3%A9lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_2017>

<https://fr.wikipedia.org/wiki/Liste_de_sondages_sur_l%27%C3%A9lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_2012>

<https://fr.wikipedia.org/wiki/Liste_de_sondages_sur_l%27%C3%A9lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_2007>

<https://fr.wikipedia.org/wiki/Liste_de_sondages_sur_l%27%C3%A9lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_2002>

<https://fr.wikipedia.org/wiki/%C3%89lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_1995>

<https://fr.wikipedia.org/wiki/%C3%89lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_1988>

<https://fr.wikipedia.org/wiki/%C3%89lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_1981>

<https://fr.wikipedia.org/wiki/%C3%89lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_1974>

<https://fr.wikipedia.org/wiki/%C3%89lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_1969>

<https://fr.wikipedia.org/wiki/%C3%89lection_pr%C3%A9sidentielle_fran%C3%A7aise_de_1965>

**8°) Raw Forecasts’ margin of error and Forecast confidence Intervals**



**9°) (Raw) Forecast Confidence intervals and error bars for individual equations**













1. To which we add the “diverse-left.” [↑](#endnote-ref-1)